A Transgenic Analysis of the Role of Smad4 in the Mammary Gland

Eleanor K. Duff

PhD

University of Edinburgh

2000

Declaration

This thesis and the research described herein is solely my own work. Any collaborative work or assistance from others is explicitly acknowledged at the relevant point within the text.

Eleanor K. Duff, December 2000

Thesis Index Page No.

Abstract 8

Acknowledgements 10

Abbreviations 12

Chapter 1 – Introduction 16

1.1 The Transforming Growth Factor (TGF-β) Superfamily 16

1.1.1 Biological Actions of TGF-β 16

(a) TGF-β and the Immune System 18
(b) TGF-β and Wound Healing 20
(c) TGF-β and Development 22

1.1.2 Signalling Receptors 28
1.1.3 Receptor Activation 31
1.1.4 Latency of TGF-β 33
1.1.5 Propagation of the Signal 33

1.2 Smads as Mediators of TGF-β signalling 36

1.2.1 Smad Families and Function 37
1.2.2 Smad Structure 39
1.2.3 Signalling through Smads 41

1.3 Smads in the Nucleus 43

1.3.1 DNA binding of Smads 45
1.3.2 DNA Binding Partners for the Smads 45
1.3.3 Activation of Transcription 48
1.3.4 Repression of Transcription 49
1.3.5 Modulation of Transcription 50

1.4 TGF-β and the Mammary Gland 51

1.4.1 Mammary Gland Structure and Function 52
1.4.2 Smads in the Mammary Gland 57

1.5 Animal Models of Gene Function 58

1.5.1 Gene Targeting in ES cells 59
1.5.2 Cre-Lox Technology 63

1.6 Aims of this Project 66

Chapter 2 - Materials and Methods 68

2.1 Preparation of DNA 68

2.1.1 Large-scale preparation of plasmid DNA – Maxiprep 68
2.1.2 Small-scale preparation of plasmid DNA – Miniprep 69

2.2 Manipulation of DNA 70

2.2.1 Restriction Endonuclease Digestion of Plasmid DNA 70
2.2.2 Agarose Gel Electrophoresis of DNA 71
2.2.3 Phenol: Chloroform Extraction and Ethanol Precipitation of DNA 71
2.2.4 Estimation of DNA Concentration and Purity 71
2.2.5 Isolation of DNA fragments from Agarose gels 72
2.2.6 Modification of the ends of Linear DNA: Dephosphorylation 72
2.2.7 Klenow Reaction 72
2.2.8 T4 DNA Polymerase 73
2.2.9 Southern Blot Hybridisation 73
2.2.10 Dideoxynucleotide Chain Termination Sequencing of Plasmid DNA 74

2.3 Targeting Construct Preparation 75

2.3.1 Screening the λ PS mouse genomic library 75
2.3.2 Creating the floxed Smad4-HPRT targeting vector 77
2.3.3 Constructing the BLG-Smad4 transgene 78
2.3.4 Preparation of Transgene for Pronuclear injection 78
2.3.5 Collection of Fertilised Oocytes and pronuclear injection 79
2.3.6 Extraction of DNA from Mouse Tails 80
2.3.7 PCR protocol 80
2.3.8 DNA Digestion and Southern Transfer 81
2.3.9 Labelling DNA probes and Hybridisation of Southern Membranes 81

2.4 General Cell Culture Techniques (Teviot Place Protocols) 82

2.4.1 Maintenance of Embryonic Stem (ES) cells 82
2.4.2 Defrosting ES cells 82
2.4.3 Gelatinisation of Tissue Culture Plates 82
2.4.4 Complete Media (CM) 83
2.4.5 Trypsinisation of ES Cells 83
2.4.6 Freezing ES Cells 83
2.4.7 Freeze Media (FM) 84
2.4.8 Electroporation of ES Cells 84
2.4.9 G418 Selection and Picking of ES Cell Clones 85
2.4.10 Thymidine kinase and Selection in Ganciclovir and Picking of ES cell clones 86
2.4.11 Extraction of ES Cell DNA to Enable Subsequent Genetic Analysis 86
2.4.12 Lysis Buffer 87

2.5 General Culture of ES cells (NPU Protocols) – Differences in protocols which may affect the outcome of gene targeting 87

2.5.1 Splitting ES Cells 87
2.5.2 NPU Protocol LIF Complete Medium 88
2.5.3 NPU Protocol In Vitro Cre Targeting 88
2.5.4 LIF Complete Medium + G418 (250μg/ml) 89
2.5.5 Ganciclovir 90
2.5.6 LIF Complete Medium + 2μM Ganciclovir 90
2.5.7 Freezing Cells 90
2.5.8 First Round Targeting of ES cells with a NEOMYCIN cassette 91
2.5.9 Karyotyping ES cells 92
2.5.10 10x HBS (Hepes-Phosphate Buffered Saline) 93

2.6 Animal Breeding 94

2.6.1 Floxed Stat3 mice 94
2.6.2 BLG-Smad4 Transgenic mice 94
2.6.3 Harvesting Mammary Glands 94

2.7 Mammary Gland analysis 95

2.7.1 Immunoblotting 95
2.7.2 Immunoprecipitation 96
2.7.3 Electrophoretic Mobility Shift Assay 96
2.7.4 Histological analysis 97
2.7.5 Immunohistochemistry 97

2.8 Transfection of ES cells and further analyses **98**

2.8.1 Transfection of ES cells *98*
2.8.2 Annexin V Analysis of transfected ES cells *99*
2.8.3 Vindelov Analysis of transfected ES cells *99*

Chapter 3 - *In vitro* analysis of Smad proteins **100**

3.1 Introduction **100**

3.1.1 TGF-β and Apoptosis *100*
3.1.2 Smads and Apoptosis *100*

3.2 Results **101**

3.2.1 Smad Overexpression constructs *101*
3.2.2 Embryonic Stem Cells *104*
3.2.3 Annexin V analysis after over-expression of Smads in ES cells *105*
3.2.4 Vindelov Analysis after over-expression of Smads in ES cells *110*

3.3 Discussion **111**

Chapter 4 - Generation and Characterisation of a BLG-Smad4 Transgenic line **115**

4.1 Introduction **115**

4.1.1 Transgenic Mouse Models *115*
4.1.2 Mouse Models – the Pros and Cons *115*
4.1.3 The Transgenic Approach *117*
4.1.4 The Promoter *117*
4.1.5 Other Considerations *119*
4.1.6 The Beta-Lactoglobulin Promoter (BLG) *120*

4.2 Results **121**

4.2.1 Making the Smad4 Transgene *121*
4.2.2 Testing Founder Mice *125*
4.2.3 Analysis of the Mammary Gland Smad4 Transgenic Mouse Model *130*

4.3 Discussion 137

Chapter 5 - Generation of a Conditional Knockout of Smad4 in the Mammary Gland 140

5.1 Introduction 140

5.1.1 Gene Targeting 140
5.1.2 Cre Lox Technology 141

(a) The Cre-Lox reaction 142
(b) The LoxP sequence 144

5.1.3 Applications Of Cre Lox Technology 144

5.2 Results 146

5.2.1 Screening The λ Library 146
5.2.2 Creating The Floxed Smad4-Hprt Targeting Vector 151
5.2.3 Generation Of Correctly Targeted Embryonic Stem (ES) Cells 157

5.3 Teviot Place Analysis 157

5.3.1 Southern Blot Analysis Of First Round Targeting 157
5.3.2 Administration Of Cre Recombinase To Remove The Floxed Neo/TK 160
5.3.3Cre-mediated deletion of the Floxed Neo/YK Cassette 163
5.3.4 Karyotyping Correctly Targeted ES Cell Clones 166

5.4 NPU Analysis 167

5.4.1 Targeting The Smad4 Locus At The NPU 167

5.5 Discussion 169

5.5.1 Floxed Smad4 mice – where next? 169

Chapter 6 - Smad-STAT Interactions in the Mammary Gland 174

6.1 Introduction 174

6.1.1 Mammary Gland Involution 174
6.1.2 Smads and STATs and Apoptosis 176
6.1.3 Co-Activators and Complexes 179

6.2 Results **181**

6.2.1 Smads are up regulated at the start of involution 181
6.2.2 Smad protein levels in a model of delayed involution, a conditional STAT3 knockout 187
6.2.3 PAI-1 and p21/WAF1/Cip1 levels in the mammary gland 187
6.2.4 DNA Binding Activity of Smads in the Mammary Gland 189

6.3 Discussion **191**

Chapter 7 – Discussion **199**

7.1 Introduction **199**

7.2 Smads and Apoptosis **199**

7.3 Transgenic Animal Models **200**

7.3.1 Creation of the BLG-Smad4 Transgenic 200
7.3.2 Future Work with the BLG-Smad4 Transgenic 202
7.3.3 Creating the Conditional allele of Smad4 204
7.3.4 Future work with the Conditional "Floxed" Smad4 Model 204

7.4 Smad-STAT Interactions in the Mammary Gland **205**

7.5 Conclusions **205**

Bibliography **207**

Appendix A: Solutions and Distributors **234**

Appendix B: Publications **239**

Abstract

Mammary gland involution proceeds through massive, highly controlled epithelial cell apoptosis and tissue remodelling. Thus the mammary gland represents an ideal physiological environment in which to study apoptosis. Recently the Smad gene family have been identified as mediators of TGF-β superfamily signalling and have been implicated in mediating epithelial cell apoptosis. The Smad family of signal transduction proteins focuses around a central mediator, Smad4, and this thesis reports a number of analyses to investigate the potential apoptotic role of this and other Smad proteins.

To investigate the effects of Smad proteins in normal euploid cells, various Smads were over-expressed in embryonic stem cells (ES cells). Analysis of ligand-specific Smads 1, 2, 3 and the central mediator Smad4, showed that only Smad4 could induce both a G1 arrest and apoptosis. This confirmed a role for Smad 4 in engaging apoptosis, and these studies were then extended to an *in vivo* model of apoptosis, namely involution of the mammary gland. This was undertaken using two transgenic approaches. First, a mammary-specific Smad4 transgenic strain was created which utilised the ovine beta-lactoglobulin gene (BLG) promoter to drive Smad4 expression to the mammary epithelium during lactation. Remarkably, mice transgenic for Smad4 and characterised as over expressing Smad4 throughout lactation and involution showed no overt phenotype. This result may arise because levels of Smad4 are already limiting in the normal mammary gland, because functional activity of Smad4 is being repressed, or even because Smad4 does not mediate a central role in the apoptotic response in the mammary gland. The second transgenic approach was to create mice with a conditional Smad4 allele using the Cre-Lox system. A floxed targeting construct was successfully used to target ES cells and these were then subsequently used to create chimeras. Germline transmission of the floxed Smad4 allele and subsequent crossing to a BLG-Cre transgenic will generate a powerful tool to investigate the requirement for Smad4 specifically within the mammary epithelium. Finally, Smad-STAT interactions were analysed biochemically during involution of the mammary gland. Other workers had previously demonstrated these

pathways interact *in vitro* and this was confirmed *in vivo* with data suggesting a novel role for STAT3 as an inhibitor of Smad activity during involution.

In conclusion this thesis has addressed the role of the Smad proteins, and especially the role of the central mediator Smad4, in mediating apoptosis. I have shown that Smad4 can mediate apoptosis in normal euploid cells in culture, but that over expression of Smad4 in the mammary epithelium apparently fails to alter the programme of involution. This result may arise as a consequence of interaction with other pathways, as I have also demonstrated crosstalk between the Smad and STAT signalling pathways *in vivo*. Finally, I have used Cre-Lox technology to create a conditional Smad4 allele, which can now be used to facilitate detailed, tissue specific analysis of Smad4 function *in vivo*.

Acknowledgements

This project was funded by the Wellcome Trust as part of Edinburgh University's first 4-year PhD. Throughout the 4 years the Trust has supplied infinite amounts of information, training, courses and (most importantly!) money and for this I gratefully acknowledge their support. I would also like to thank my supervisors Prof Alan Clarke, Prof Martin Hooper, Dr John Mason and Dr Stefan Selbert who have always been helpful, encouraging and inspirational - without their vast knowledge and guidance it would all undoubtedly have gone pear-shaped.

The following people have been an essential part of my life here at the department for a number of reasons. I would like to thank them individually for their time, expertise, advice and general loveliness: Dorothy Watson for organising the first Edinburgh Wellcome 4 year PhD rabble (us); Jonathan Bard for worldly-wise words; Sarah Howie for explaining the intricacies of the Flow cytometer; Rachel Chapman for mentoring and many discussions on MG maintenance; Liz Lovejoy for Vector Nti help and sticky buns; Jean Manson, Nadia Tuzi and Jennifer Doig for help at the NPU; "Germline Joe" Mee and Austin Smith for blastocyst injections at the CGR; Roberta James and Alasdair MacKenzie for oocyte injections; Norah Spears & lab for coffee, biscuits and a lot of fun; John Verth, Linda Clark and Ian McCall for help with the EKs and Bob Morris and Frances Rae for making sure that I never set my hair on fire more than once. Imaging has been a large part of this project and could not have been attempted without the help and advice of the imaging gurus Stuart McKenzie and Steven Haig.

Everyone in the department has been helpful and encouraging throughout my PhD but sometimes it takes a special bunch of folk and a good pint or two to sort things out. For this, and more, I am indebted to Jan Cullingworth, Jason Read, Jon Coe, Owen Sansom, Chris Armit and Martyn Link. Last, but not least, I would like to thank Kevin for all his support over the years and mum, for everything.

"Mammaries, light the corners of my mind…"

Abbreviations

Abd-A	Abdominal A
ActR-I	Activin receptor I
ActR-II	Activin receptor II
AdCre	Adenovirus Cre
ALK	Activin Like Kinase
AMH	Anti-Müllerian Hormone
AP1	Activator Protein-1
ATF2	Activated Transcription Factor-2
BFGF	basic Fibroblast Growth Factor
BHLH	basic Helix-Loop-Helix
BLG	β-Lactoglobulin
BMP	Bone Morphogenic Protein
BMPR-I	Bone Morphogenic Receptor I
BMPR-II	Bone Morphogenic Receptor II
CBP	CREB Binding Protein
CNS	Central Nervous System
Co-Smad	Collaborating Smad
CRE	cAMP Response Element
CSF-1	Colony Stimulating Factor-1
Dad	Daughters against Dpp
DAF	Dauer Larva Formation
DMSO	Dimethyl sulphoxide
DNA	Deoxyribonucleic Acid
DPC4	Deleted in Pancreatic Carcinoma-4
Dpp	Decapentaplegic
EAE	Experimental Autoimmune Encephalitis
ECM	Extracellular Matrix
EDTA	Ethylenediaminetetraacetic acid
EGF	Epidermal Growth Factor

ER	Oestrogen Receptor
ES cells	Embryonic Stem Cells
FAST	Forkhead Activin Signal Transducer
FGF	Fibroblast Growth Factor
FKBP-*	Forkhead Binding Protein
GDF-9	Growth and Differentiation Factor-9
GDNF	Glial Derived Neurotrophic Factor
GH	Growth Hormone
H&E	Haematoxylin and Eosin
HAT	Histone Acetyl Transferase[1]
HDAC	Histone De-Acetylase
HEPES	N-2-Hydroxyethyl piperazine-N'-2-ethanesulfonic acid
Hh	Hedgehog
HSV	Herpes Simplex Virus
IFN-γ	Interferon-γ
IGF-1	Insulin like Growth Factor-1
IL-*	Interleukin *
I-Smad	Inhibitory Smad
IV	Intravenous
JAK	Janus Activated Kinase
LAP	Latency Associated Peptide
LIF	Leukaemia Inhibitory Factor
LoxP	Locus of crossover of P1
MAD	Mothers Against Dpp
MADR	Mad-Related
MAPK	Mitogen Activated Protein Kinase
MBP	Myelin Basic Protein
MH1	Mad Homology 1
MH2	Mad Homology 2
MHC	Major Histocompatibility Complex

* denotes a number as in IL-2, IL-6 etc

MIS	Müllerian Inhibiting Substance
MMP	Matrix Metalloprotease
MOPS	3-[N-Morpholino]propanesulphonic acid
MS	Multiple Sclerosis
NFκB	Nuclear Factor κ B
NK	Natural Killer
NO	Nitric Oxide
OT	Oxytocin
PA	Plasminogen Activator
PAI-1	Plasminogen Activator Inhibitor-1
PC5	Proprotein Convertase 5
PCR	Polymerase Chain Reaction
PDGF	Platelet Derived Growth Factor
PEBP2/CBF	Polyoma virus Enhancer Binding Protein2/ Core Binding Factor
PECAM-1	Platelet Endothelial Cell Adhesion Molecule-1
polβ	Polymerase β
PR	Progesterone Receptor
R-Smad	Receptor activated Smad
SARA	Smad Anchor for Receptor Activation
Sax	Saxophone
SBE	Smad Binding Elements
SIP-1	Smad Interacting Protein-1
Ski	Sloan Kettering avian retrovirus gene
Sma	Small
Smad	Sma-mad
Smurf1	Smad Ubiquitination Regulatory Factor
SnoN	Ski-related Novel gene (N= non alu containing)
STAT	Signal Transducer and Activator of Transcription
TFE3	Transcription Factor E3
TGF-β	Transforming Growth Factor-β

TGF-α	Transforming Growth Factor-α
TGIF	TGF Interacting Factor
Th1	T helper cell type 1
Th2	T helper cell type 2
TIMP	Tissue Inhibitor of Metalloproteases
TK	Thymidine Kinase
Tkv	Thickveins
TNF-α	Tumour Necrosis Factor-α
TR-I	TGF Receptor I
TR-II	TGF Receptor II
TRIP-1	TGF-β Receptor Interacting Protein-1
TRIZMA	Tris(hydroxymethyl)aminomethane
SDS	Sodium dodecyl sulphate
Ubx	Ultrabithorax
VDRE	Vitamin D Response Element
Wg	Wingless
X-gal	5-Bromo-4-Chloro-3-Indolyl-ß-D-Galactopyranoside

Chapter 1 - Introduction

1.1 The Transforming Growth Factor-β (TGF-β) Superfamily

The TGF-β superfamily comprises an extremely large number of structurally related polypeptide growth factors, each capable of regulating a fascinating array of cellular processes, including cellular proliferation, lineage determination, differentiation, motility, adhesion and death. Expressed in complex temporal and tissue-specific patterns, TGF-β and related factors play a prominent role in the development, homeostasis and repair of virtually all tissues in organisms from fruitfly to human.

TGF-β and related factors are multifunctional agonists whose effects depend on the state of responsiveness of the target cells as much as on the factors themselves. Given this varied nature it is not surprising in retrospect that the gradual discovery of these factors over the past years has been made through very disparate lines of investigation. For example, the founding member of the family, TGF-β1, was identified as a regulator of mesenchymal growth (Heine *et al*, 1987) and, separately, as an antimitogen in epithelial cells (Kurokowa, *et al*, 1987). The same is true of other members of the family, for example Activins and Bone Morphogenic Proteins (BMPs), and enforces the idea that members of this family can have varied effects dependent on the situation or circumstance. This phenomenon of "multi-tasking" ligands had remained unexplained for more than fifteen years until the discovery of the Smad family of signal transduction/transcription molecules in 1996. However, before this point in time much had been elucidated about the ligands themselves and activation of their receptor families, and also about target genes which encode immediate effectors of cellular responses, such as cell cycle regulators and extracellular matrix components.

1.1.1 Biological Actions of TGF-β

TGF-β and its related factors control cell fate by controlling proliferation, differentiation and apoptosis and are therefore important for both the development and maintenance of nearly all tissues (Hogan, 1996). TGF-β has a multitude of actions, the most important of which may be to regulate genes whose products contribute to the growth and stability of the extracellular matrix (ECM). Thus it stimulates the production of collagen, fibronectin and other ECM components and decreases the production of proteases that degrade the ECM. Since these early discoveries, TGF-β has been demonstrated to be both a potent negative regulator of growth for a wide variety of cell types, including cells derived from epithelial, endothelial, neuronal, haematopoietic and lymphoid origins and a potent mitogen for other cell types, for example mesenchymal cells. The growth inhibitory effect of TGF-β is thought to be the molecular basis for many more of the global effects of TGF-β including the regulation of immune responses, development and cellular differentiation. In addition to its effects on cell proliferation and differentiation, TGF-β also regulates many biological processes through its ability to transcriptionally regulate a huge number of important cellular genes (Massague, 1990). For example, TGF-β regulates wound healing, cellular adhesion and extracellular matrix deposition through its ability to co-ordinately control the expression of extracellular matrix proteins and the proteases which degrade them.

TGF-β stimulates the anchorage-dependent proliferation of some cells and inhibits the proliferation of others. Although the ability of TGF-β to affect different cell types in opposite ways is puzzling, it may not reflect fundamental differences in the initial cellular responses to TGF-β. Instead, the different types of cellular responses may arise because TGF-β initiates a number of changes in all responsive cells, some of which may lead to proliferation and others to proliferative arrest. Depending on the individual responses of specific cell types and on the environment of the cells, the balance of the effects of these changes could lead to cellular proliferation or inhibition of proliferation. TGF-β is a cytokine commonly associated with

differentiation of cells. For example in mammals it is found at highest concentrations in the tip of the intestinal villus where it is associated with the most differentiated cells (Murphy, 1998). This observation has given rise to the postulate that it is involved in an "autocrine axis" that regulates the orderly progression of differentiated cells not only in the intestine but also in the skin, perhaps explaining its vital role in wound healing (Wahl *et al*, 1989). Other members of the family, especially the Bone Morphogenic Proteins, are also very important in the healing processes that occur in bone and cartilage (Bostrom and Asnis, 1998).

(a) TGF-β and the Immune System

TGF-β is produced by a wide range of immune cells including macrophages, NK cells, T cells and B cells, and has both pro- and anti-inflammatory properties, depending on its environment and concentration (reviewed in Wahl, 1994). Perhaps the most important pro-inflammatory property of TGF-β is its ability to recruit monocytes, T cells and neutrophils to the site of inflammation early in an infection, through modulation of endothelial cell adhesion molecule expression. However, at high concentrations TGF-β can have strong anti-inflammatory properties including suppression of TNF-α and NO from macrophages (Alleva et al, 1994), inhibition of IFN-γ and TNF-α from NK cells (Bellone *et al*, 1995) and it can also antagonise IFN-γ-stimulated up regulation of major histocompatibility complex (MHC) class II antigens (Nandan and Reiner, 1987). Antagonism of IL-4, IL-2 (Reugemer *et al*, 1990) and IL-12 by TGF-β is also well documented (Pardoux *et al*, 1997). The importance of TGF-β in limiting inflammation is clearly demonstrated in TGF-β-knockout mice, which die at 3–4 weeks of age from a wasting syndrome associated with multifocal inflammatory disease (Shull *et al*, 1992). Pathological examination of null mice has revealed an excessive inflammatory response with massive infiltration of lymphocytes and macrophages in many organs, but primarily in heart and lungs. Many lesions in this model resemble those found in autoimmune disorders, graft-vs. -host disease or in certain viral diseases (Kulkarni *et al*, 1993 (a) and (b)). This mouse model is also characterised by a host of other symptoms which

confirm the multifunctional nature of TGF-β1. Evidence is constantly accumulating which shows that TGF-β is a key cytokine which regulates immune function *in vivo,* and that overproduction of TGF-β may be associated with immunosuppression. Cytokines are known to be important mediators of immune regulation by regulatory cells, and similarly the direct administration of cytokines can affect autoimmune processes and influence the recovery from autoimmune disease. It has been shown by Santambrogio *et al.* (1993) that the administration of TGF-β provides protection against autoimmune encephalomyelitis (EAE), a condition which is thought to reflect certain aspects of multiple sclerosis (MS) in humans, and can be induced in rodents by injection of myelin antigens such as myelin basic protein (MBP) in complete Freund's adjuvant. T-cell lines and clones that transfer EAE in the rat and mouse produce the Th1 cytokines IFN-γ, IL-2, tumour necrosis factor-α (TNF-α) and TNF-β, and these cytokines are also present in the CNS of animals with active disease. These authors found that injection of TGF-β on days 5–9 after immunisation was highly protective and therefore concluded that the protective effect of TGF-β was exerted at the level of the target organ and/or its vascular endothelium rather than through a direct effect on lymphoid cells.

It thus appears that Th2-type cytokines like TGF-β are important regulators of immune responses. Experiments (Rocken *et al*, 1994) have shown that the induction of EAE in mice by the transfer of the Th1-like or pro-inflammatory cells can be inhibited by the administration of IL-4 or TGF-β at the time of cell transfer and this and other similar findings have been the basis for the current trend in gene therapy experiments which have used myelin basic protein- specific T cells engineered to express latent transforming growth factor-β in an experimental treatment of murine EAE (Chen *et al*, 1998).

(b) TGF-β and Wound Healing

One of the most exciting recent TGF-β discoveries is that TGF-β enhances wound healing. When painted on a wound TGF-β has been found to increase both the strength of the newly formed connective tissue and its rate of formation. This effect is complex, involving both fibroblasts and macrophages and both are stimulated to migrate to the site of a wound. Fibroblasts increase their synthesis of collagen and fibronectin and decrease their secretion of proteases that tend to degrade the extracellular matrix. Stimulation of secretion of the fibroblast mitogen, IL-1, by monocytes has also been observed. The source of TGF-β at wounds may be platelets which release TGF-β when they degranulate. In general, the release and activation of TGF-β stimulates the production of various extracellular matrix proteins and inhibits the degradation of these matrix proteins, although exceptions to these principles abound. These actions of TGF-β contribute to tissue repair, which under ideal circumstances leads to the restoration of normal tissue architecture and may involve a component of tissue fibrosis. In many diseases, excessive TGF-β contributes to a pathologic excess of tissue fibrosis that compromises normal organ function, a topic that has been the subject of numerous reviews (Roberts and Sporn, 1989) but which will be discussed succinctly here.

TGF-β is secreted by a variety of cells and serves multiple functions in tissue/organ repair by increasing cellular production of ECM components, such as fibronectin and collagen, and cellular expression of matrix receptor integrins. Furthermore, the synthesis of Plasminogen Activator Inhibitor-1 (PAI-1) and Tissue Inhibitor of Metalloproteases (TIMP) is also increased by TGF-β, while the expression of collagenase and Plasminogen Activator (PA) is decreased. This up regulation of inhibitor synthesis and down regulation of protease synthesis further augments the accumulation of ECM proteins induced by TGF-β, and is the basis for fibrotic tissue formation caused by the action of TGF-β. The mechanism of receptor signalling

holds the key to TGF-β regulation of cellular responses and is one of the most intensively studied areas of TGF-β research.

The use of *in vitro* systems has aided the identification and characterisation of many of the components that interact during wound healing. For instance, there are many compelling *in vitro* examples of the control of cellular gene expression through the adhesive interaction of connective tissue cells with their surrounding ECM and many of these interactions are mediated through cell adhesion receptors called integrins (Hynes, 1992). Frequently, gene expression is regulated by cytokines and growth factors, such as TGF-β, released from adjacent cells or the surrounding ECM through limited proteolysis. Indeed, proteolytic degradation of ECM is an essential feature of tissue repair and remodelling processes. The serine proteinases, including plasminogen activator (PA)—plasmin and the matrix metalloproteinases (MMPs), are the two major groups of ECM-degrading enzymes that interact and form a lytic cascade for ECM remodelling. The major function of PA is to control the activation of plasminogen into plasmin. Plasmin is not only the primary effective enzyme in fibrinolysis, but it also participates in the breakdown of other ECM proteins, and activates procollagenase into collagenase (a member of the MMP family) (DeClerck, and Laug, 1996). Thus, the initiation of the proteinase cascade by PA leads to a notable amplification of proteolytic activity. The complexity of this regulatory system is increased by the fact that plasmin can release active TGF-β from its latency-associated protein (Lyons *et al*, 1990). In turn, TGF-β regulates plasminogen activator inhibitor 1 (PAI-1), MMPs, tissue inhibitor of metalloproteinases 1 (TIMP-1) and genes encoding ECM components and their integrin receptors (Roberts and Sporn, 1990). TGF-β is the archetypal example of a growth factor involved in this process because it is acknowledged to have the broadest range of activities, in both normal and fibrotic repair injury. Many other growth factors such as epidermal growth factor (EGF), fibroblast growth factor (FGF) and platelet-derived growth factor (PDGF) also have substantial influence on the growth and differentiation of keratinocytes, fibroblasts and endothelial cells during wound repair.

Mechanisms of foetal wound repair have attracted much attention in recent years, owing to the observation that foetal wounds heal without scar formation. Foetal skin contains abundant hyaluronic acid and does not become inflamed during repair. Nevertheless, a key to foetal scarless healing seems to reside in the expression of TGF-β. Whereas adult wound sites contain high and persistent levels of TGF-β1, TGF-β1 is expressed transiently at low levels in foetal wounds (Martin *et al*, 1993). An understanding of the mechanism of TGF-β expression and the mode of action on cell proliferation/differentiation and ECM production/remodelling may ultimately allow an understanding of the control of excess scarring in adult wounds.

(c) TGF-β and Development

Expression patterns of TGF-βs during embryogenesis as well as the activities of these molecules in *in vitro* assays of biological processes relating to development have suggested that TGF-βs play an important role in embryonic development. Embryonic induction, the process by which signals from one cell population influence the fate of another, plays an essential role in the development of all organisms so far studied. In many cases, the signalling molecules belong to large families of highly conserved proteins, originally identified as mammalian growth factors, the largest of which is the TGF-β family. Genetic studies in *Drosophila* on the TGF-β related gene, *decapentaplegic* (*dpp*), have revealed the existence of conserved mechanisms regulating both the expression of the protein during development and the way in which it interacts with other signalling molecules to generate patterns within embryonic tissues.

Recently, several groups have published exciting new data on the function of *dpp* in *Drosophila*. These studies raise the possibility that *dpp* regulates and is regulated by two other secreted growth factors, encoded by *wingless* (*wg*) and *hedgehog* (*hh*). As both *hh* and *wg* -type (*Wnt*) genes exist in vertebrates, work on *dpp* in *Drosophila* provides an important paradigm for the understanding of TGF-β related gene function in higher organisms.

The *dpp* gene is necessary for several different morphogenetic processes during *Drosophila* embryogenesis, including dorsal/ventral patterning of the embryo, establishment of the proximo-distal axis of appendages, eye development, and midgut morphogenesis (Spencer *et al*, 1982, Irish and Gelbart 1987, Panganiban *et al*, 1990). In dorsal/ventral patterning, evidence suggests that the DPP protein acts as a morphogen regulating dorsal cell fate (Ferguson and Anderson, 1992). Wharton *et al*. (1993) have demonstrated that increasing the number of *dpp* genes in either mutant or wild-type embryos results in the formation of an increasing number of dorsal type cells in a dose-dependent manner.

During midgut development, DPP is made in the visceral mesoderm and induces changes in gene expression in the adjacent endoderm, thus providing a model for understanding epithelial/mesenchymal interactions in higher organisms. Studies have shown that *dpp* transcription in the mesoderm is activated by the homeobox gene *Ultrabithorax* (*Ubx*) and repressed by *Abdominal-A* (*Abd-A*) (Manak *et al*, 1994). In both cases, regulation is direct and involves multiple DNA-binding sites in a 5' upstream region of *dpp* (Capovilla *et al*, 1994 and Masucci and Hoffman, 1993). In the endoderm, DPP switches on the homeobox gene *labial,* which has a DPP-response element in its 5' upstream region (Tremml and Bienz, 1992). These studies clearly show that *dpp* acts during development both upstream and downstream of homeotic genes.

Studies on eye morphogenesis demonstrate another key role for *dpp* during *Drosophila* development (Heberlein *et al*, 1993). In this organ, *dpp* is expressed in the morphogenetic furrow, a wave of organising activity that moves across the eye imaginal disc. The extracellular signalling molecule encoded by *hh* regulates transcription of *dpp*, and DPP induces morphogenesis in front of the furrow (Ma *et al*, 1993). In the leg disc, *dpp* is also regulated by *hh* (Basler and Struhl, 1994) and cooperates with *wg* in the regulation of the homeobox gene *aristaless*, which then acts as an organiser for proximal/distal patterning (Campbell *et al*, 1993). In this

case, at least three signalling molecules, *hh*, *dpp*, and *wg*, coordinate to generate patterning in an epithelial sheet.

Several other members of the TGF-β family are also involved in early patterning of the embryo: Vg1 and members of the *Xenopus* nodal-related (Xnr) family (Lustig *et al*, 1996) can induce dorsal mesoderm formation, the first step towards neuralisation. In contrast, BMP4 acts as a mesoderm ventraliser (Jones *et al*, 1996); mutation of cleavage sites yields BMP molecules that are incapable of maturation and leads to dorsalisation of the early *Xenopus* embryo (Hawley *et al*, 1995).

These experiments support the concept that neuralisation of embryonic cells, formerly attributed to a positive action of the Spemann organizer, occurs through a default pathway opposed by members of the TGF-β family (Green 1994, Hemmati-Brivanlou and Melton, 1997). Three articles published in the same issue of Cell provide evidence that the *Xenopus* embryo derived Spemann organizer products Noggin and Chordin interrupt Dpp and BMP signalling in *Xenopus* (Zimmerman *et al*, 1996 and Piccolo *et al*, 1996) and also in *Drosophila* (Holley *et al*, 1996) by binding Dpp or BMP-4, and so preventing them from reaching their respective receptors. In keeping with this hypothesis, Noggin does not antagonise the effect of constitutively active forms of the Dpp receptors (Holley *et al*, 1996).

These experiments also highlight the conservation of patterning molecules of the TGF-β family between species; despite the large evolutionary distance between insects and vertebrates, both Dpp and BMP4 act alike in *Xenopus* and *Drosophila* (Jones *et al*, 1996). TGF-β family members also influence patterning in fish: a constitutive form of a novel type I receptor is able to induce the most anterior dorsal mesoderm in the zebrafish (Renucci *et al*, 1996).

During mouse embryogenesis, BMPs and their receptors are essential to mesoderm development (Mishina *et al*, 1995). One of the most exciting findings in the field of mouse embryogenesis is that Nodal is expressed asymmetrically in the mouse

embryo and may be involved in the establishment of the left–right axis (Collignon *et al*, 1996).

Three TGF-β family members — inhibin, MIS (also known as anti-Müllerian hormone or AMH) and GDF-9 — are preferentially, if not exclusively, expressed in the gonads. All three happen to be relatively remote members of the family, with only 20–30% identity to the family ancestors Dpp/BMP and to TGF-β itself. MIS is responsible for the regression of Müllerian ducts in male mammalian foetuses and also in chick embryos (Behringer, 1994) and Neeper *et al*, 1996). As with all TGF-β family members, MIS requires cleavage at a proteolytic site to exert biological activity. Work from the laboratory of Holly Ingraham (Nachtigal *et al*, 1996) suggests that the processing enzyme could be the kex2/subtilisin-like endoprotease PC5. Co-expression of PC5 and MIS in transfected mammalian cells results in efficient processing and PC5 is present in foetal testes, coincident with the beginning of MIS expression.

Gene knockout experiments have revealed that MIS synergises with inhibin to influence gonadal tumour development. Inactivation of MIS (Behringer *et al*, 1994) or the genes encoding the MIS receptor (Mishina *et al*, 1996) leads to Leydig cell hyperplasia. Male and female mice in which the inhibin genes have been inactivated by gene targeting develop granulosa/Sertoli cell tumours, first detectable at around 4 weeks of age, followed by a wasting syndrome caused by activin signalling (Coerver *et al*, 1996). Mice deficient in inhibins and either MIS (Matzuk et *al*, 1995) or the MIS receptor developed testicular tumours in both the intratubular and interstitial compartments earlier. Inhibin-deficient mice with concurrent inactivation of gonadotropin-releasing hormone do not develop tumours at all (Kumar *et al*, 1996).

Mothers against Dpp (MAD) and proteins related thereto, have recently emerged as key players in the signal transduction field. The *Drosophila* MAD protein is required for any response to Dpp. Null alleles of MAD suppress the dominant phenotype of constitutively active mutants of *thick veins* (*tkv*), which encodes a DPP type I

receptor, proving that MAD is an essential component of the downstream signalling pathway of Dpp (Hoodless *et al*, 1996). Homologues of the *Drosophila* MAD protein have been cloned in *C. elegans* (Morita *et al*, 1999), *Xenopus* (Graff et *al*, 1996), mouse (Baker and Harland, 1996) and human (Hoodless *et al*, 1996, Zhang *et al*, 1996 and Riggins *et al*, 1996).

A more distantly related protein, DPC4, which associates with mammalian MAD-related proteins (MADR) proteins, is mutated in pancreatic cancers (Hahn *et al*, 1996). MADR proteins appear to be specific for a given member of the TGF-β family, suggesting that they are dedicated to transducing the signals for a specific subclass of TGF-β ligands (Thomsen 1996). For instance, MADR2 (Smad2) is phosphorylated by TGF-β (Eppert *et al*, 1996) and activin (Baker and Harland, 1996) and not by BMPs. In *Xenopus*, MADR1 (Smad1) specifies a ventral phenotype similar to the one produced by BMP2 whereas MADR2 (Smad2) induces a dorsal phenotype observed in TGF-β and activin signalling. In *C. elegans*, inactivation of the MAD-related gene *sma* yields phenotypes similar to the one produced by the mutation of DAF-4, which encodes a BMP receptor.

Drosophila is not the only system providing genetic clues to the biological importance of TGF-β related genes. Important contributions are now beginning to come from mice with mutations or targeted disruptions of genes in this family. The short ear (se) mutant has deletions or rearrangements in the gene for BMP-5 that is associated with loss of specific skeletal structures, indicating a role for the gene product in skeletogenesis (Kingsley *et al*, 1992). In the case of the mouse mutation 413-d, studies have shown that it results from a retroviral insertion in the TGF-β related gene *nodal*. During gastrulation, homozygous mutants are unable to form axial structures, including organized dorsal mesoderm and notochord (Zhuo *et al*, 1993, Conlon *et al*, 1991, Varlet *et al*, 1997). Targeted disruptions of TGF-β1, inhibin, and activin/inhibin aB genes all result in aberrant phenotypes in homozygous mutant mice. Mice that are homozygous null mutant for TGF-β1 are viable for a few weeks before dying from a 'wasting' syndrome (Shull *et al*, 1992). These results

indicate that the embryo does not need to produce TGF-β1 for prenatal development, although the mother may provide some TGF-β1 protein in the uterus and milk. Mice deficient for inhibin are also viable, but develop gonadal stromal tumours, suggesting that one function of the protein is to regulate or suppress stromal cell proliferation in this organ (Matzuk *et al*, 1992). Targeted disruption of the related gene encoding activin/inhibin aB has also been reported (Vassalli *et al*, 1994). Homozygous null mutants are viable, but females have reduced fertility, indicating a role for activin/inhibin aB in reproductive function. However, the fact that some viable offspring are born to these females eliminates activin aB from playing a role in embryonic pattern formation.

Comparative studies on BMP-4 in *Xenopus* and mouse point to a conserved role in specifying posteroventral mesoderm during gastrulation (Hogan *et al*, 1994). TGF-β knockout mice can again confirm this essential role as mice lacking TGF-beta 3 exhibit an incompletely penetrant failure of the palatal shelves to fuse leading to cleft palate (Proetzel *et al*, 1995). The fact that TGF-β1 knockout animals are viable and present with an immunological phenotype and not a severe developmental phenotype has been explained in two separate ways as follows; 50% of the TGF-β1 conceptuses die at mid-gestation from defective yolk sac vasculogenesis, a severe early developmental defect. The other 50% are developmentally normal but die three weeks postpartum from massive inflammation. It has been suggested that dichotomy in TGF-β1 lethal phenotypes is due to either maternal TGF-β1 rescue of some, but not all, TGF-β null embryos (Letterio *et al*, 1994) or is simply dependant on the genetic background of the conceptus (Bonyadi *et al*, 1997).

In summary TGF-β appears to be a peptide that "gives form to things", e.g., it regulates the activities of other growth factors and may even be able to organise cells into functional units.

1.1.2 Signalling Receptors

Protein kinases play an important role in signal transduction by phosphorylating specific amino acids of downstream substrates. Catalytic domains predict what kind of residues will be phosphorylated by a given kinase. Many receptors for hormones and growth factors are membrane-bound tyrosine kinases with reasonably well-understood signal transduction pathways. In contrast, until relatively recently serine/threonine kinase activity had been detected only in cytoplasmic proteins downstream of receptor signalling. Starting with the *Caenorhabditis elegans* DAF-1, a myriad of genes encoding transmembrane serine/threonine kinases have been cloned; almost all act as receptors for members of the TGF-β family. As discussed above, members of the TGF-β family are found in species ranging from *Drosophila* to human and can be grouped in small clusters progressively diverging from *decapentaplegic* (*Dpp*) and its human homologue, bone morphogenetic protein-2 (BMP-2). With the exception of glial cell line derived neurotrophic factor (GDNF), all members of the TGF-β family in all species signal through a receptor complex formed by two distantly related types of serine/threonine kinase proteins. Christened type II and type I on the basis of their molecular weights, two such receptors (TR-II and TR-I) have now been functionally characterised (Wrana *et al*, 1994). Recombinant type II receptors bind their cognate ligand on their own, whereas type I receptors do so only when co-expressed with an appropriate type II receptor. Dpp/BMP receptors exhibit somewhat different binding properties, insofar that type II receptors bind ligand efficiently only in the presence of type I, which, in turn, can bind free ligand (albeit weakly) in the absence of type II (Yamashita *et al*, 1996). The kinase activity of TR-II is constitutive and its phosphorylation state is not modified by ligand binding (Wrana *et al*, 1994). Ligand binding induces the formation of a receptor complex, most likely a heterotetramer containing two molecules each of TR-I and TR-II (Weis-Garcia and Massagué, 1996, Luo and Lodish, 1996, Ten Dijke *et al*, 1996). The central event in ligand-induced TGF-β receptor activation is the transphosphorylation of TR-I by TR-II at threonines and serines located in the GS box, a conserved glycine/serine-rich domain located immediately upstream of the

kinase consensus domains (Franzen *et al*, 1995). Absence of TR-I, or mutations thereof leading to either loss of kinase activity or inability to be transphosphorylated by TR-II, blocks the signalling response (Weis-Garcia and Massagué, 1996). A missense mutation of TR-II that blocks the recognition of TR-I as a substrate has the same effect (Carcamo *et al*, 1995). In contrast, a mutation within the GS box of several type I receptors yields a constitutively active molecule that signals in the absence of ligand and receptor II (Wieser *et al*, 1995). Truncation of TR-II immediately after the transmembrane domain produces dominant negative molecules which bind to ligand and type I receptors but do not induce biological responses (reviewed in Brand and Schneider, 1996). The model for activation of the TGF-β receptor complex has recently been extended to activin. Although it contains a transmembrane domain, the type I receptor acts as a substrate and signal transducer for TR-II, the primary receptor (Massagué and Weis-Garcia, 1996). The composition of the receptor complex, and more particularly the identity of the type I receptor, determines the nature of the signal. Establishing which type I receptor associates with which ligand/type II receptor complex is difficult because a given ligand may signal through different receptor complexes and a given receptor can recognize different ligands.

This family was initially discovered in 1991 with the cloning of the first activin receptor (Mathews and Vale, 1991) now called ActR-II. Based on their structural and functional properties, the TGF-β receptor family is divided into two subfamilies; Type I receptors and Type II receptors and there are currently 12 members of the Type I subgroup and 7 members of the Type II subgroup (**Figure 1.1**). Vertebrate Type I receptors form three groups whose members have similar signalling activities. In mammals these groups include 1) TβR-1, ActR-1B and ALK7, 2) BMPR-1A, BMPR-1B and 3) ALK1 and ALK7. In invertebrates Type I receptors include Thick veins (Tkv) and Saxophone (Sax) which both act as receptors for Dpp in *Drosophila*, and *C. elegans* has Daf-1, a Type I receptor for the BMP-like ligand Daf-7. The type II receptor subfamily in vertebrates includes TβR-II and BMPR-II, which bind TGF-β and BMPs respectively, and the ActR-II and ActR-IIB which bind activins.

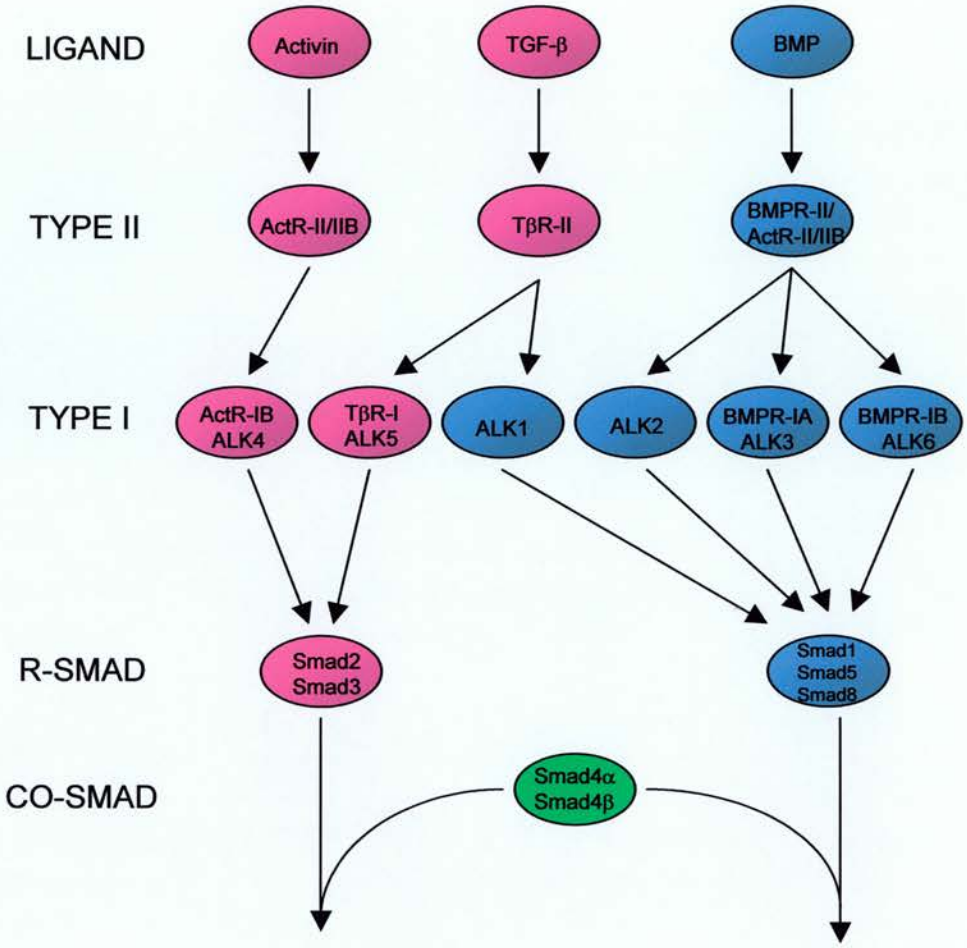


Figure 1.1.

Vertebrate TGF-β Receptors

The Transforming Growth Factors-β (TGF-β), activins and bone morphogenic proteins (BMPs) signal through distinct sets of Type I and Type II receptors and Smad proteins. Five distinct Type II receptors and seven Type I receptors, also termed Activin-like kinases (ALKs) have been identified. Type I receptors act downstream of Type II receptors

Both the Type I and Type II receptors are glycoproteins of approximately 55kDa and 70kDa respectively with a relatively short extracellular domain and no singular structural features in the transmembrane region. However, a unique feature of the Type I receptors is a highly conserved domain preceding the protein kinase domain called the GS domain. This region is so called due to the fact that it contains a characteristic SGSGSG sequence and it has been proved to be a key regulatory region that controls the catalytic activity of the Type I receptor kinase (Wrana *et al*, 1994). The kinase domain in both receptors is a typical serine/threonine kinase domain (Matthews and Vale, 1991). Type I receptors have been shown to phosphorylate their substrates, the Smads, on serine residues, while Type II receptors phosphorylate themselves and Type I receptors on serine and threonine residues.

1.1.3 Receptor Activation

TGF-β and related family members activate signalling by binding to and bringing together pairs of Type I and Type II receptors. Two general modes of binding ligand have been documented. One mode, characteristic of TGF-β and activin receptors, involves direct binding to the Type II receptor and subsequent interaction of this complex with the Type I receptor, which then becomes recruited to the complex. Here Type I receptors for these factors can recognise ligand that is bound to the Type II receptors but not ligand that is free in solution. This mode of binding was discovered using TGF-β-resistant cell mutants – TGF-β1 can bind to TβR-II in cell mutants lacking TβR-I but cannot bind to TβRI in mutants lacking TβR-II (Laiho *et al*, 1990, Boyd and Massague 1989). The second mode of binding is typical of BMP receptors and is co-operative, involving Type I and Type II receptors that bind ligand with high affinity when expressed together but low affinity when expressed separately (Liu *et al*, 1995).

The kinase activity of TβR-II is constitutive and its phosphorylation state is not modified by ligand binding (Wrana *et al*, 1994) Ligand binding induces the formation of a receptor complex, most likely a heterotetramer containing two

molecules each of TβR-I and TβR-II (Weis-Garcia *et al*, 1996, Luo and Lodish, 1996 and Ten Dijke *et al*, 1996) The central event in ligand-induced TGF-β receptor activation is the transphosphorylation of TβR-I by TβR-II at threonines and serines located in the GS box (Franzen *et al*, 1995). Absence of TβR-I, or mutations thereof lead to either loss of kinase activity or an inability to be transphosphorylated by TβR-II which then blocks signalling responses (Weis-Garcia and Massague, 1996). A missense mutation of TβR-II that blocks the recognition of TβR-I as a substrate has the same effect (Carcamo *et al*, 1995). In contrast, a mutation within the GS box of several type I receptors yields a constitutively active molecule that signals in the absence of ligand and receptor II (Wieser *et al*, 1995). Truncation of TβR-II immediately after the transmembrane domain produces dominant negative molecules which bind to ligand and type I receptors but do not induce biological responses (reviewed in Brand and Schneider 1996). The model for activation of the TGF-β receptor complex has been extended to activin (Willis *et al*, 1996, Attisano *et al*, 1996 and De Winter *et al*, 1996) and although it contains a transmembrane domain, the type I receptor acts as a substrate and signal transducer for TβR-II; the primary receptor (Massague and Weis-Garcia 1996). The composition of the receptor complex, and more particularly the identity of the type I receptor, determines the nature of the signal.

Furthermore, given that type I and II receptors which probably do not associate under normal cellular conditions may do so when over expressed in cell lines, binding data alone are not sufficient to predict the existence of a functional signalling complex. Demonstration of co-immunoprecipitation is more reliable as it requires high-affinity association (Vivien and Wrana 1995); proof of the capacity of the receptor under investigation to restore signalling in mutant cell lines is even more conclusive (ten Dijke et *al*, 1994). The ultimate test, however, is the examination of the phenotypic effects of inactivating the receptor during normal development. In man, mutations have been described only for the anti-Müllerian hormone (AMH)/Müllerian Inhibiting Substance (MIS) type II receptor (Imbeaud *et al*, 1995) and for the activin receptor-like kinase-1 (ALK-1) (Berg *et al*, 1997). In the absence of natural

mutations, the generation of mutated receptor molecules has proven very useful for dissecting the signalling pathway for TGF-β family members, bone morphogenic factors and their receptors in body patterning.

1.1.4 Latency of TGF-β

Like other members of the family, TGF-β is synthesised as the C-terminal domain of a precursor form that is cleaved before secretion from the cell (Derynck *et al*, 1985). However, the TGF-β propeptide, which is referred to as the latency associated peptide or LAP, remains non-covalently bound to TGF-β after secretion, retaining TGF-β in a latent form that cannot bind the receptors. Most cell types secrete TGF-β in this biologically inert form (Roberts and Sporn 1990). *In vitro* TGF-β may be activated by acid, alkali, heat or limited proteolysis (Miyazono *et al*, 1993). In tissue culture activation of latent TGF-β may involve a combination of steps including LAP proteolysis and cell-cell interactions, but the real physiological activation mechanism remains to be defined.

1.1.5 Propagation of the Signal

Recent TGF-β receptor phosphorylation studies by Wrana *et al*. (Wrana *et al*, 1994) have provided insights into the molecular mechanisms by which TGF-β receptors are activated. TβR-II is a constitutively active kinase, the autophosphorylation of which is not clearly up regulated upon ligand binding. After binding of TGF-β to TβR-II, TβR-I is recruited into a stable complex and is phosphorylated by TβR-II on its serine and threonine residues, which are predominantly located in the GS domain, the 30 amino acid region preceding the kinase domain and conserved in type I receptors for other TGF-β-related factors which contains SGSGSG repeats. (Wrana *et al*, 1994). Phosphorylation of TβR-I is important for signal transduction, as impaired responses are seen in cells with a mutation in TβR-II that renders its kinase inactive (Wrana *et al*, 1992), or mutations of serine and threonine residues in the GS domain of TβR-I (Wieser *et al*, 1995), or a mutation in TβR-II that inactivates its

ability to recognize TβR-I as a substrate (Cárcamo *et al*, 1995). Activated TβR-I does not autophosphorylate or phosphorylate TβR-II (Wrana *et al*, 1994), but it is phosphorylated by cytoplasmic kinases, albeit to a lower extent than is TβR-II (Weiser *et al* 1995). Recently, a constitutively active TβR-I mutant (in which the threonine residue at position 204 was replaced by an aspartic acid residue) was described, and was shown to have an increased *in vitro* kinase activity and to signal in the absence of ligand and TβR-II (Weiser *et al*, 1995). Thus, TβR-I functions as a substrate for TβR-II; its phosphorylation and activation by TβR-II is essential and sufficient for most TGF-β mediated signalling.

Using a yeast two-hybrid interaction screen with an intracellular type I receptor as bait, FKBP-12 (Wang *et al*, 1996), a binding protein for FK506 and rapamycin, and farnesyl protein transferase (Kawabata *et al*, 1995), an enzyme subunit involved in farnesylation or geranylgeranylation of proteins such as Ras, Rac and Rho, were isolated. Both proteins failed to interact with a kinase-defective type I receptor; however, until recently their involvement in TGF-β signalling remained to be established as no ligand dependence in receptor association or phosphorylation had been demonstrated. FKBP12 is now known to interact with Type I receptors and inhibits the phosphorylation of the Type I receptors by the Type II receptors (Wang *et al*, 1996). Using the TβR-II intracellular domain as bait, TβR interacting protein-1 (TRIP-1) was isolated (Chen *et al*, 1995). TRIP-1 is homologous to WD-domain-containing proteins that are involved in various protein–protein interactions. The WD repeat comprises a 44–60-residue sequence within many proteins although the exact significance of this structure is unknown. The interaction and phosphorylation of TRIP-1 is dependent on TβR-II kinase activity, but is independent of ligand and TβR-I. The functional significance of TRIP-1 in TGF-β signalling remains to be determined.

Drosophila *dpp* signals via *dpp* type I receptors, *thick veins* and *saxophone*, and an activin/dpp type II receptor, *punt* (Nellen *et al*, 1994, Penton *et al*, 1994, Brummel *et al*, 1994, Letsou *et al*, 1995, Ruberte *et al*, 1995). Loss of *saxophone* or *thick veins* activity results in phenotypes which resemble partial and complete loss of *dpp* gene function, respectively (Nellen *et al*, 1994); *thick veins* may be required for all *dpp*-specified cell fates, whereas *saxophone* may modulate responses at high concentrations of *dpp*. Mutations that abolish *punt* activity show a similar phenotype to that obtained with complete loss of *dpp* gene function. *Thick veins* and *punt* are strictly required for *dpp* signalling, and their *in vivo* loss of function cannot be compensated for by other receptors (Ruberte *et al*, 1995). A downstream component in the *dpp*-signalling pathway, *schnurri*, was recently found to encode a putative zinc-finger-containing transcription factor that regulates the expression of *dpp*-responsive genes (Arora *et al*, 1995, Grieder *et al*, 1995). Further genetic screens in *Drosophila* will probably lead to the identification of additional genes that encode components downstream of *dpp* receptors; mammalian homologues of these components may act downstream of TGF-β, activin and BMP receptors. Moreover, inactivating mutations in *daf-1* and *daf-4*, which encode *Caenorhabditis elegans* serine/threonine kinase receptors, have been shown to lead to constitutive dauer larvae development; genetically identified *C. elegans* genes that act downstream of these receptors may have mammalian counterparts that participate in intracellular signalling of TGF-β superfamily members (Estevez *et al*, 1993).

Signals emanate from a TGF-β Type I receptor when it is phosphorylated by its activator, the Type II receptor. Ligand binding induces the formation of a heteromeric complex of Type I and Type II receptors (Franzen *et al*, 1993) and because the ligands are dimeric in nature each might contact one Type I and Type II receptor resulting in a heterotetrameric complex. This is supported by data from 2-D gel electrophoresis (Yamashita *et al*, 1994) and co-precipitation of receptor studies (Weiss-Garcia and Massague, 1996). Formation of this ligand-induced complex rapidly leads to phosphorylation of the Type I receptor (with TGF-β and activin receptors) and this reaction is catalysed by the Type II receptor. Until recently this marked the end of the known pathway of TGF-β signal transduction as substrates for

the Type I receptor were unknown. A role of the Type I receptor as the downstream signalling component in the receptor complex was originally inferred from the observation that the kinase activity of TβRI is required for signal transduction and yet its substrate is neither TβR-I nor TβR-II (Wrana *et al*, 1994). The proteins of the Smad family are the first identified substrates of Type I receptor kinases and play a pivotal role in the transduction of receptor signals to target genes in the nucleus.

1.2 Smads as Mediators of TGF-β Signalling

The first identified Smad protein was the product of the *Drosophila Mad* gene (*mothers against dpp*). *Mad* was identified in 1996 as a maternally expressed enhancer of the *dpp* phenotype in a screen for mutations that would exacerbate the effect of weak *dpp* alleles (Raftery *et al*, 1995). Since that initial discovery many related genes have been found in nematodes, *Drosophila* and vertebrates. Mad homologues have been discovered in *C. elegans* and called *sma-2, sma-3* and *sma-4* because mutation in these genes causes small body size (Krishna *et al*, 1999 and Morita *et al*, 1999). Homologues in vertebrates have been discovered through screening EST (expressed sequence tag) databases and cDNA libraries. These proteins have been named Smads (for *sma/mad* related) and the family currently encompasses 10 members (for review see Ten Dijke *et al*, 2000).

Initial evidence that Smads function downstream of TGF-β receptors was the observation that mutant *Drosophila Mad* genes could inhibit signalling by a hyperactive *thick veins* (*Tkv*) receptor. The most compelling evidence came from the observation that in response to TGF-β and related family members, Smads are phosphorylated (Hoodless *et al*, 1996, Zhang *et al*, 1996), accumulate in the nucleus (Hoodless *et al*, 1996, Baker and Harland 1996) and become transcriptionally active (Liu *et al*, 1996). This evidence placed the Smads directly downstream of TGF-β receptors.

1.2.1 Smad Families and Functions

Based on structural and functional considerations, vertebrate Smads fall into three subfamilies; receptor-regulated Smads (R-Smads), the Co-Smad Smad4, and the inhibitory Smads (I-Smads), Smad6 and Smad7 (for review see ten Dijke *et al*, 2000). The family is highly conserved and homologues of each of these classes have been identified in *Xenopus*, *Drosophila* and *C. elegans* (**Figure 1.2**). The R-Smads are activated and phosphorylated only in response to specific ligands and this phenomenon is important in maintaining specificity in TGF-β signalling. Thus, the TGF-β and activin Type I receptors only activate Smad2 and Smad3, whereas the BMP Type I receptors all target Smad1, Smad5 and Smad8 (Kretzschmar *et al*, 1997). Signalling by R-Smads requires the participation of a collaborating Smad or Co-Smad. There is currently only one vertebrate Co-Smad, Smad4, although others have been determined in *Xenopus* (Howell *et al*, 1999). Although Smad4 is similar to the R-Smads in overall structure, it is not normally phosphorylated in response to agonists. Smad4 is generally required for Smad2 or Smad3-dependent growth inhibitory responses in mammalian cells, and a dominant-negative Smad4 construct interferes with Smad1 and Smad2 signalling in frog embryos and mammalian cells (Zhang *et al*, 1996 and Lagna *et al*, 1996). Smad4 therefore participates in TGF-β, activin and BMP signalling pathways as a shared partner of receptor-regulated Smads. The *Drosophila Medea* and *C. elegans sma-4* gene products are close homologues and probably fulfil a similar function in these organisms.

In contrast to this the I-Smads function to block TGF-β and BMP signalling. These Smads form stable complexes with the activated receptors and inhibit signalling by preventing access and phosphorylation of the R-Smads. Vertebrate Smad7 inhibits both BMP and TGF-β receptors whereas Smad6 appears to be more specific to the BMP pathway only. *Drosophila Dad* inhibits *Mad* signalling (Tsuneizumi *et al*, 1997). Transcription of Smad6 and Smad7 is induced by TGF-β, activin and BMP, providing a negative feedback mechanism for regulation of Smad signalling.

1.2.2 Smad Structure

Smad proteins contain highly conserved N-terminal and C-terminal domains referred to as MH1 and MH2 domains respectively and a middle linker region of variable length and sequence which is proline-rich (de Caestecker *et al*, 1997). The MH2 domain contains SSxS motifs which are involved in receptor phosphorylation and also mediate a large number of distinct protein-protein interactions which include: association with the membrane-bound receptor (Macias-Silva *et al*, 1996), assembly of heteromeric R-Smad-Smad4 complexes (Hata *et al*, 1997) and the association of Smads with DNA binding proteins and transcriptional co-activators (Liu *et al*, 1997 and Chen *et al*, 1997). The MH1 domain has been suggested to auto inhibit the MH2 domain (Kim *et al*, 1997) and also has been shown to be capable of binding DNA directly (Song *et al*, 1998 and Dennler *et al*, 1998).

The crystal structure of the MH2 domain of Smad4 has been analysed and found to contain five α helices (H1 to H5) and three loops (L1 to L3) (Shi *et al*, 1997). This structure has revealed that many of the mutations in Smad proteins identified in cancers map to the protein-protein interface between R-Smads and Smad4 (Shi *et al*, 1997). Functional analysis of R-Smads has identified the L3 loop and H1 region as the important determinants controlling specificity over the R-Smad-receptor interaction at the cell surface (Lo *et al*, 1998). A similar study with the Type I receptor has shown that L45 region in the kinase domain specifies interactions with the Smads (Chen *et al*, 1998, Feng and Derynck, 1997), thus Smad-receptor interactions may be mediated by L3 on the Smad and L45 on the receptor (**Figure 1.3**).

The crystal structure of the MH1 domain of Smad3 has been analysed in the presence of a short DNA fragment containing two of the known Smad binding sites called Smad Binding Elements (SBE) (Shi *et al*, 1998). These elements consist of 5'-GTCT-3' repeats and have been found in the promoters of various genes known to be activated by Smads/TGF-β. The MH1 domain is composed of four α-helices, six

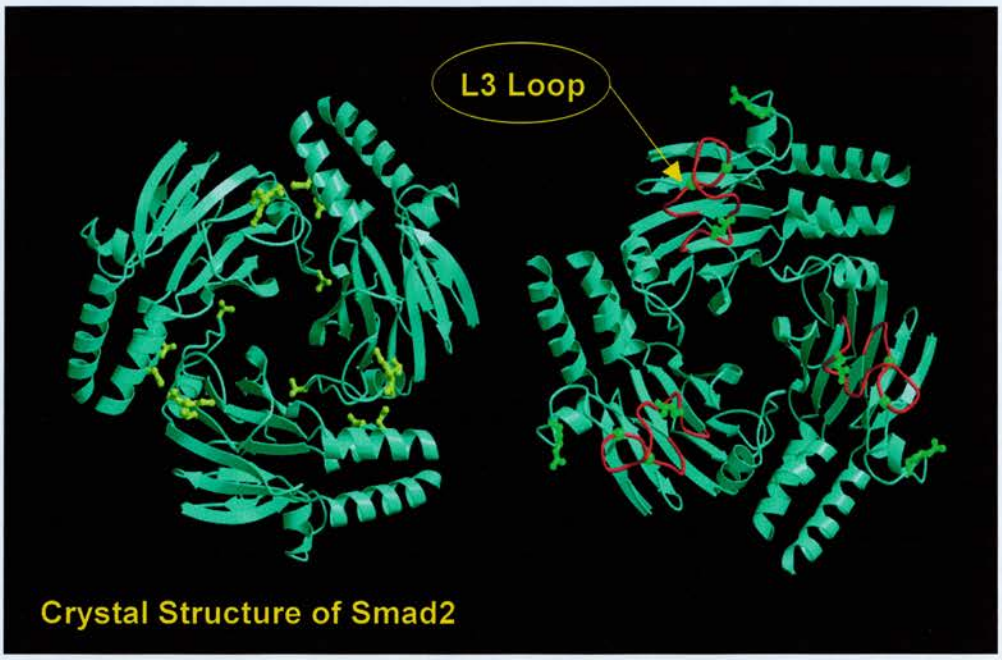


Figure 1.3

Crystal Structure of Smad2

The figure shows a stereo image of a Smad2 trimer. The crystal structure of a number of Smad proteins has recently been elucidated and various groups have identified a discrete surface structure in Smads 1 and 2 that mediates and specifies their receptor interactions. This structure is the L3 loop, a 17 amino acid region that protrudes from the core of the conserved SMAD C-terminal domain. On the figure above the L3 loop is represented in red on the MH2 domains of Smad2. These studies have identified the L3 loop as a determinant of specific SMAD-receptor interactions, and indicate that the L3 loop, together with the C-terminal tail, specifies SMAD activation.

short β strands and five loops and DNA binding is accomplished by a structure composed of a β hairpin that contacts DNA in the major groove (Shi *et al*, 1998). The linker region that connects the MH1 and MH2 domains is much less conserved, but it also contains important peptide motifs which can control Smad activity. These include a conserved proline rich PY motif that may be important for crosstalk with the MAPK pathway (Kretschmar *et al*, 1997, 1999).

1.2.3 Signalling through Smads

In the basal state, Smads exist as homo-oligomers that reside in the cytoplasm. Upon ligand activation of the receptor complex, the Type I receptor kinase phosphorylates specific Smads, which then form a complex with Smad4 and move into the nucleus. In the nucleus, these complexes either alone or in association with other proteins activate target genes by binding to specific promoter elements.

Smads are serine-phosphorylated in response to agonists – Smad1 in response to BMP2 or 4 (Hoodless *et al*, 1996), Smad2 in response to TGF-β or activin (Eppert *et al*, 1996) and Smad3 in response to TGF-β (Zhang *et al*, 1996). Although the kinetics of this phosphorylation are relatively slow when transfected Smads are used ($t_{½}$ ~ 5 minutes) Smads have been shown to be direct substrates of the receptors. Smad1 is phosphorylated by a highly purified BMPR-I kinase domain (Kretschmar *et al*, 1998), Smad2 by immunoprecipitated TGF-β receptor complexes (Macias-Silva *et al*, 1996) and Smad3 by TβR-I kinase preparation (Zhang *et al*, 1996). *In vitro* and *in vivo* receptor mediated phosphorylation occurs at serines in the C-terminal motif SSxS of Smad1 and Smad2. This motif is also present in the other R-Smads of all the vertebrate and non-vertebrate Smads, but not the Co-Smads or I-Smads. This is in agreement with the observation that Smads4, 6 and 7 are not phosphorylated in response to ligand binding. Mutation of the serines in this sequence inhibits receptor-mediated phosphorylation of Smads1 and 2 *in vivo* and *in vitro* and their association with Smad4, accumulation in the nucleus and mediation of transcriptional responses

(Kretschmar *et al*, 1997), proving that phosphorylation of this motif is required for Smad activation.

Interactions between receptors and Smads are a critical step in initiating the intracellular signalling cascade. Recently a protein named SARA (Smad Anchor for Receptor Activation) was identified in a screen for Smad2 interacting proteins and has been shown to play a pivotal role in the association of the Smad with the receptor (Tsukazai *et al*, 1998). SARA specifically binds unphosphorylated Smad2 and Smad3 and contains a FYVE domain adjacent to its Smad-binding domain. The FYVE domain is a double zinc-finger motif that in other proteins has been shown to bind phosphatidylinositol-3-phosphate and can thereby anchor proteins to membranes. In SARA, this FYVE domain is required for its proper subcellular localisation into discrete areas. Furthermore SARA recruits Smad2 to these areas where activation of TGF-β signalling causes Smad2 to dissociate from SARA. Interestingly, these subcellular domains also contain the TGF-β receptors and SARA can also associate with heteromeric TGF-β receptor complexes through a carboxy-terminal region that is distinct from the Smad binding domain. Essentially this data has shown that SARA can recruit Smads to the membranes and regions of the cell where the activating receptor kinase is localised. The elucidated crystal structure of the Smad2 MH2 domain bound by SARA has shown that the MH2 domain interacts with an extended proline rich motif in the Smad binding domain of SARA. An important asparagine residue (N381) has been identified and substitution of this residue with a serine results in interference in the binding of Smad2 to SARA, a failure of SARA to localise the Smad protein and loss of activity from reporter genes (Wu *et al*, 2000). Thus SARA plays an important role in the initiation of TGF-β signalling by controlling the localisation of the Smad substrate.

Smad2 and Smad3 become transiently and selectively associated with the activated TGF-β receptor complex and the phosphorylation sites in the MH2 region of the Smad become phosphorylated. This then facilitates Smad2 dissociation from the receptor as shown by phosphorylation defective Smad2 mutants which have

enhanced interactions with the receptors (Macias-Silva *et al*, 1996). This transient nature of the Smad-receptor interaction is consistent with the role of Smads as carriers of receptor signals into the nucleus.

Receptor-phosphorylated Smads associate with Smad4, which functions as a shared partner required for transcriptional activation. Based on structural considerations and the observation that mutations in the Smad4 L3 loop abolish the ability of Smad4 to associate with Smad2, the L3 loop in Smad4 appears to mediate the association with receptor activated Smads (Shi *et al*, 1997).

Nuclear translocation of receptor activated Smads occurs with kinetics that closely follow those of the ligand-induced phosphorylation and association with Smad4. Nuclear translocation of Smads1 and 2 does not always require Smad4, as determined using Smad4-defective cells and Smad4 is also translocated into the nucleus in response to TGF-β or BMP; as this translocation requires the presence of Smad1 or Smad2 it appears that receptor activated Smads bind Smad4 in the cytoplasm and carry it into the nucleus (Liu *et al*, 1997) (**Figure 1.4**).

1.3 Smads in the Nucleus

Initial studies on the function of Smads as transcriptional regulators have revealed that Smads can bind directly to DNA. However, more recent work showing the interactions of Smads with a diverse array of DNA-binding proteins have suggested that the primary role of Smads is not to target specific genes through their DNA binding activity but rather to function as co-modulators of transcriptional activity.

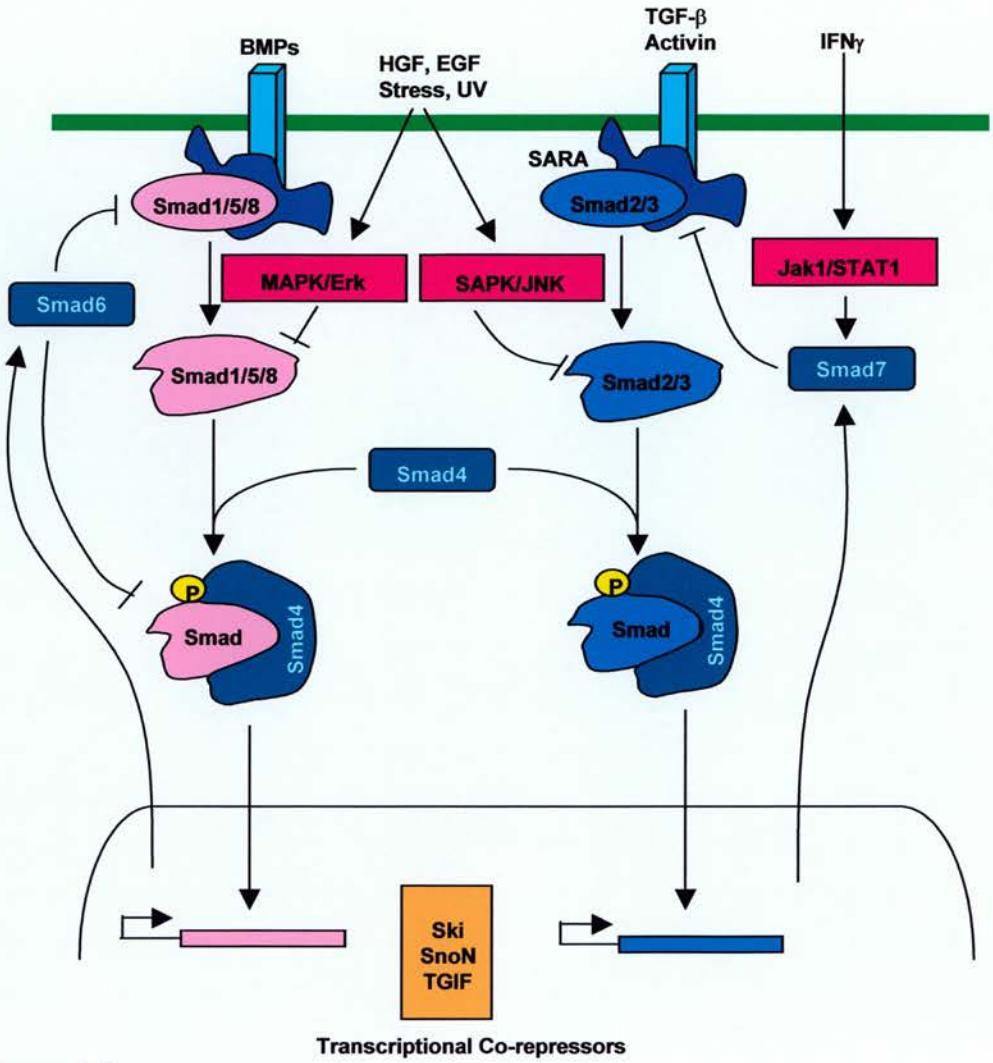


Figure 1.4

The Smad Signalling Pathway

The Smad signalling pathway gives specificity to the myriad responses of TGF-β/activin and bone morphogenic proteins bound at the cell surface. Ligand binds at the cell surface inducing the formation of heteromeric complexes of specific type I and type II serine/threonine kinase receptors. The type I receptor is then phosphorylated and activated by the type II receptor and this leads to activation of specific Smad proteins. Smads 1,5 and 8 are activated by BMPs and Smads2 and 3 are activated by TGF-β and activins. These R-Smads then associate with a Co-Smad (in this case Smad4) and the complex then translocates to the nucleus where together with other proteins they direct transcriptional responses. A number of negative regulators of this pathway exist including the I-Smads Smad6 and Smad7 and transcriptional co-repressors Ski, SnoN and TGIF.

1.3.1 DNA Binding of Smads

In the nucleus, R-Smad-Co-Smad complexes are involved in transcriptional regulation of target genes (Massague *et al*, 1998). The MH1 domains of Smad3 and Smad4 were found to possess an intrinsic property to bind to specific DNA sequences that contain 5'-AGAC-3' sequences, termed Smad-binding elements (SBEs) (Yingling *et al*, 1997, Zawel *et al*, 1998, Dennler *et al*, 1998). Direct binding of Smads to DNA was first described for Drosophila MAD which binds a GC-rich element in the DPP-responsive *vestigial* gene – mutation of this site prevented *dpp*-dependent expression of a reporter gene *in vivo* (Kim *et al*, 1997). Extensive studies with Smad3 and Smad4 similarly established that these Smads bind DNA and, in most cases, mutation of Smad binding sites in TGF-β-responsive promoter elements reduces activation of reporter constructs (for review see Derynck *et al*, 1998). These studies suggest that Smad DNA binding is important for activating target genes, however Smads appear to bind DNA with rather low specificity and affinity. Selection of Smad3 and Smad4 binding sequences yielded a consensus-binding site of two inverted repeats of GTCT (Zawel *et al*, 1998). The crystal structure of Smad3 bound to this sequence shows that MH1 monomers bind only to a single copy of this core sequence (Shi *et al*, 1998). Further analysis has suggested that a GNCN repeat is sufficient for Smad-DNA binding (Johnson *et al*, 1999). As this occurs frequently in promoters, it suggests that Smad-DNA binding alone is unlikely to be sufficient for Smad-dependent targeting of specific genes. Furthermore, Smad2 has a unique insert in the MH1 domain that abolishes DNA binding (Yagi *et al*, 1999). These findings suggest that interactions with DNA-binding partners are essential for Smads to function in regulating specific transcriptional responses.

1.3.2 DNA Binding Partners for the Smads

Assuming that Smad-DNA binding alone is not sufficient for gene activation, there are now a growing number of examples in which Smads have been found to co-operate with DNA binding partners to regulate transcription. The most extensively characterised of these is the FAST family of DNA binding proteins. *Xenopus* FAST

is a winged helix/forkhead DNA-binding factor that binds to an activin responsive element in the promoter region of the *Mix.2* gene (Chen *et al*, 1996, Chen *et al*, 1997). Human and mouse homologues of FAST (FAST-1 and FAST-2) have been identified (Labbé *et al*, 1998, Weisberg *et al*, 1998, Zhuo *et al*, 1998, Liu *et al*, 1999). In addition to regulating *Mix.2*, mouse FAST has been shown to bind a TGF-β activin responsive region of the *goosecoid* promoter (Labbé *et al*, 1998). In the absence of signalling, FAST binds constitutively to these response elements, but is unable to activate transcription. However, in the presence of TGF-β or activin a complex containing FAST, Smad2 and Smad4 assembles on the DNA and transcription is strongly activated (see previous references and Yeo *et al*, 1999). Analysis of the complex shows that Smad2 interacts directly with FAST to recruit Smad4 into the nuclear complex. The Smad4 MH1 domain then binds DNA at an adjacent Smad binding element and this contact is essential to stabilise the DNA binding by the Smad-FAST complex.

Smads can also co-operate with DNA-binding proteins which are involved in mediating signals for other pathways. Several TGF-β responsive elements have been found to contain AP1 binding sites which are activated by heterodimers of c-jun and c-fos. Over-expression of Smad3 strongly activates several AP1-containing promoters and Smad3 can interact directly with c-jun via the MH1 domain and indirectly with c-fos via c-jun (Zhang *et al*, 1998). In the absence of these two proteins TGF-β and Smad3 only weakly induce AP1 elements (Zhang *et al*, 1998, Liberati *et al*, 1999) thus suggesting that Smads require active c-fos/c-jun dimers as DNA-binding partners in order to stimulate AP1-containing elements.

The transcription factor ATF2 is a basic helix-loop-helix (bHLH) leucine zipper protein which binds to cAMP response elements either as a homodimer or as a heterodimer with c-jun. Both Smad3 and Smad4 have been shown to bind ATF2 through their MH1 domains and in transient assays, co-operate with ATF2 to activate CRE-containing elements (Hocevar *et al*, 1999, Sano *et al*, 1999). Smads co-operate with activated ATF2, which is a target of p38 kinase cascade, and blocking either

ATF2 function or p38 activation also blocks the ability of Smads to regulate CRE reporter genes. Thus Smads can regulate transcription through interactions with ATF2 and like the Jun kinase cascade can crosstalk with the p38 pathway.

Another member of the bHLH-leucine zipper family of transcription factors, TFE3 binds to DNA through an E-box element. Smad3 and TFE3 together activate TGF-β induced transcription by binding to adjacent sites in the PAI-1 promoter (Hua *et al*, 1999). Interestingly, any alteration in the spacing between the Smad-binding sites and the E-box abrogates TGF-β induced transcription. Thus, proper positioning of Smad-binding sites may be an essential feature for Smad-dependent regulation of certain transcriptional targets, possibly preventing random activation of E-boxes by Smad signalling.

The association of Smads with members of the polyomavirus enhancer binding protein2/core binding factor (PEBP2/CBF) transcription factor complex was recently described (Hanai *et al*, 2000). This complex binds to active forms of Smad1, Smad2 and Smad3 and although the biological relevance of these interactions has yet to be determined it is important to note that DNA binding by both Smad3 and PEBP to their respective sites was essential for TGF-β-dependent activation of the promoter.

A similar situation has been described for vitamin D. Analysis of a vitamin D response element (VDRE) linked to a reporter gene has shown that the VDRE was responsive to TGF-β, but importantly only in the presence of vitamin D (Yanagisawa *et al*, 1999). Consistent with this Smad3 was found to act as a co-activator for the vitamin D receptor through its ability to form a complex with the steroid receptor co-activator-1 protein in a vitamin D and TGF-β-dependent manner. Again this points to Smad interactions with other pathways, in this case between the TGF-β and vitamin D pathways.

One variation in the way that Smads might act as transcriptional regulators has been described for the association between Smad1 and *Hoxc-8* (Shi *et al*, 1999). *Hoxc-8* is

a homeodomain transcription factor that may repress the osteopontin promoter and it has been shown that Smad1 expression may dislodge the inhibitory Hoxc-8 from the promoter. A similar Smad-mediated relief of transcriptional repression has been proposed for the reported interaction of Smads with Smad-interacting protein-1 (SIP-1) (Verschueren *et al*, 1999). Thus relief of repression may represent another mechanism for Smad-mediated transcriptional activation.

1.3.3 Activation of Transcription

Recent studies have shown that MH2 domains are responsible for mediating transcriptional activation. Work from a large number of groups has shown that this activation probably occurs through Smad recruitment of the transcriptional co-activators p300 and CREB binding protein (CBP) to target promoters (for review see Derynck *et al*, 1998). P300/CBP acts as a co-activator for a number of distinct transcription factors including p53, p73, NFκB and members of the STAT family. It works both by bridging transcription factors to the basal transcriptional machinery and also through its intrinsic histone acetyltransferase activity. Enhancement of Smad-mediated transcription occurs through the direct interaction of p300/CBP with the C-terminus of Smad1, Smad2 and Smad3 and is enhanced by TGF-β phosphorylation (Pearson *et al*, 1999, Nishihara *et al*, 1998 and Janknecht *et al*, 1998). While association of p300/CBP has been documented for these R-Smads at the C-terminal end of the protein, association of Smad4 with p300/CBP has shown a novel function of the middle linker region of this Smad, which interacts with the amino terminus of p300 (de Caestecker *et al*, 2000).

This interaction between Smads and p300/CBP has recently shown an interaction between two major signalling pathways – Smads and STATs. In this case p300/CBP is essential for the observed synergy between BMP and LIF in activation of GFAP, a marker of neural cells (Nakashima *et al*, 1999). The cytokines LIF (leukaemia inhibitory factor) and BMP2 signal through completely different receptors and transcription factors, namely STATs (**S**ignal **T**ransducers and **A**ctivators of

Transcription) and Smads. LIF and BMP2 were found to act collectively on primary foetal neural progenitor cells to induce astrocytes. The transcriptional coactivator p300 interacts physically with STAT3 at its amino terminus in a cytokine stimulation-independent manner, and with Smad1 at its carboxyl terminus in a cytokine stimulation-dependent manner. The formation of a complex between STAT3 and Smad1, bridged by p300, was therefore found to be involved in the cooperative signalling of LIF and BMP2 and the subsequent induction of astrocytes from neural progenitors (Nakashima *et al*, 1999).

1.3.4 Repression of Transcription

As well as activating target genes, Smads have also been shown to function as transcriptional repressors. The earliest discovery of this phenomenon was in the mouse *goosecoid* promoter. Here Smad2 was found to activate the gene in cooperation with FAST2, whereas Smad3 was unable to activate the element and actually blocked Smad2-dependent activation (Labbé *et al*, 1998). In contrast to this the activin response element of the *Xenopus Mix.2* gene appears to be activated by both Smad2 and Smad3 thus showing that Smads can act as both transcriptional activators and repressors depending on the promoter context. The mechanism behind this is currently unclear.

The recent identification of TGIF (5'TG3' interacting factor), SnoN and Ski (Sloan-Kettering avian retrovirus/ski-related gene) has also provided insights into how Smads can act as transcriptional blocks. TGIF is a homeodomain protein which when over-expressed blocks TGF-β-dependent activation of target genes (Wotton *et al*, 1999). TGIF interacts with Smad2 and appears to repress target gene expression through the recruitment of histone deacetylases to TGF-β pathway targeted promoters. Binding of TGIF occurs via the MH2 domain and is mutually exclusive with CBP/p300 interaction; the expression levels of TGIF versus CBP/p300 could thus determine the intensity of TGF-β/Smad responses in a cell by affecting the chromatin organisation through a balance between HDAC and HAT activity.

Binding of the nuclear oncoproteins Ski or SnoN to Smad2 or Smad3 represses transcription of Smad target genes and blocks the growth inhibitory effects of TGF-β (Luo *et al*, 1999, Sun Y *et al*, 1999, Stroschein SL *et al*, 1999, Akiyoshi *et al*, 1999). This block of gene expression appears to occur through the recruitment of the transcriptional co-repressor N-CoR which in turn has been described to recruit histone deacetylases to targeted genes. As these co-repressor proteins are recruited to genes that can be activated as a result of TGF-β treatment, these studies suggest that induction of TGF-β target genes may be a balance between recruitment of activators such as p300/CBP and repressors such as TGIF, Ski or SnoN. Interestingly, SnoN and, to a lesser extent, Ski are rapidly degraded through cellular proteasomes upon TGF-β treatment (Sun *et al*, 1999 and Stroschein *et al*, 1999) Thus Smads may interact with different DNA binding factors and regulate transcription both positively and negatively depending on the interacting partners.

Ubiquitin-mediated degradation of BMP R-Smads through specific interaction with Smad ubiquitination regulatory factor (Smurf1) was recently shown to be another mode by which the Smad pathway can be controlled (Zhu *et al*, 1999). Like other E3 ubiquitin ligases, Smurf1 mediates the conjugation of ubiquitin to specific target proteins. Increased expression of Smurf1 leads to a selective decrease in BMP R-Smad protein levels, thereby decreasing the cellular competence to BMP-mediated responses.

1.3.5 Modulation of Transcription

The work to date on Smad control of transcription suggests that, as the nature of these interactions appears to determine whether Smads positively or negatively regulate activation of target genes, Smads should be considered modulators of transcription. Smads may regulate transcription in two distinct ways; in the case of the FAST-regulated genes, FAST binds DNA with high affinity and specificity but is unable to activate target elements on its own. Transcriptional activation by FAST

thus appears to rely solely on Smad activity and Smads may function as the primary regulatory signal. Perhaps as FAST is only expressed during gastrulation, this mechanism may be unique in playing a critical role in Smad signalling in patterning in the early embryo.

Contrasting with this scenario, most other Smad partners identified can activate transcription independent of the Smads. These often require input from other pathways before the Smads can regulate their activity. Here Smads do not provide primary signals, but rather act as secondary signals which can modulate the outcome of these interactions.

1.4 TGF-β and the Mammary Gland

Programmed cell death (apoptosis) occurs during normal growth and development of the mammary gland. One of the most dramatic examples of apoptosis is evident during the remodelling of the breast that accompanies postlactational involution (Strange *et al*, 1995). Transgenic mouse models have demonstrated that overexpression of polypeptides such as transforming growth factor alpha (TGF-α) and insulin like growth factor I (IGF-I) can block this remodelling, suggesting that these growth factors may be acting as survival factors for the mammary epithelium (Jhappan *et al*, 1990, Weber *et al*, 1998). In contrast, transgenic mice that overexpress the growth inhibitor TGF-β show increased apoptosis in the mammary epithelium throughout mammary development, suggestive of a mechanism working to counterbalance the survival factors (Pierce *et al*, 1993, Jhappan *et al*, 1993). Experiments with mammary epithelial cell lines cultured *in vitro* have confirmed that these growth factors can indeed regulate apoptosis and survival in mammary epithelial cells; EGF, IGF-I, and basic fibroblast growth factor (bFGF) act as survival factors for mammary epithelial cells, while TGF-β induces their death (for review see Rosfjord and Dickson, 1999). In breast cancer, cytotoxic drugs and hormone ablation increase the expression of TGF-β, which may function to induce cell death by either paracrine or autocrine mechanisms (Reiss *et al*, 1997). Together, these

studies have begun to describe a complex dynamic pattern of pro- and apoptotic factors that promote the development of the mature mammary gland and that rapidly remodel the tissue after lactation.

1.4.1 Mammary Gland Structure and Function

Functional differentiation of the mammary gland is a crucial step in the reproductive cycle of mammals. The development of the gland in mice proceeds in distinct phases; in newborn mice a single primary duct is present which grows slowly during the first eight weeks of life into a ductal tree which fills the fat pad until the onset of puberty when pronounced ductal growth occurs. Development of the ducts continues in cycling virgins leading to the formation of a ductal tree which fills the entire mammary fat pad. Extensive ductal branching and alveolar growth occurs during pregnancy and is largely completed at parturition. Terminal differentiation of the alveolar epithelium is completed at the end of gestation with the onset of milk secretion at parturition. After weaning, the entire alveolar epithelium undergoes apoptosis and the gland is then remodelled. Within a few weeks the gland has the appearance of that of a mature virgin (Hennighausen and Robinson, 1998) (**Figure 1.5**).

This repetitive cycling of the gland from full lactation through involution and back again is a phenomenon that occurs in few other tissues and thus elects the mammary gland as an ideal model to study highly controlled *in vivo* cellular apoptosis. Both the role of systemic hormones and influence of the stroma on mammary epithelial cells have been recognised for some time (Sakakura, 1991). However, only now through the availability of knockout mice have we been able to dissect individual steps in the pathways of the translation of hormonal signals into morphogenetic and developmental events. Two unique aspects of mammary gland development have greatly aided in exploiting these knockout animals to elucidate the specific roles of the epithelium and the stroma. First, the mammary gland develops predominantly in the postpartum mammal. Therefore, an entire developmental program, mimicking embryonic development of other organs, can be viewed and followed in postpartum

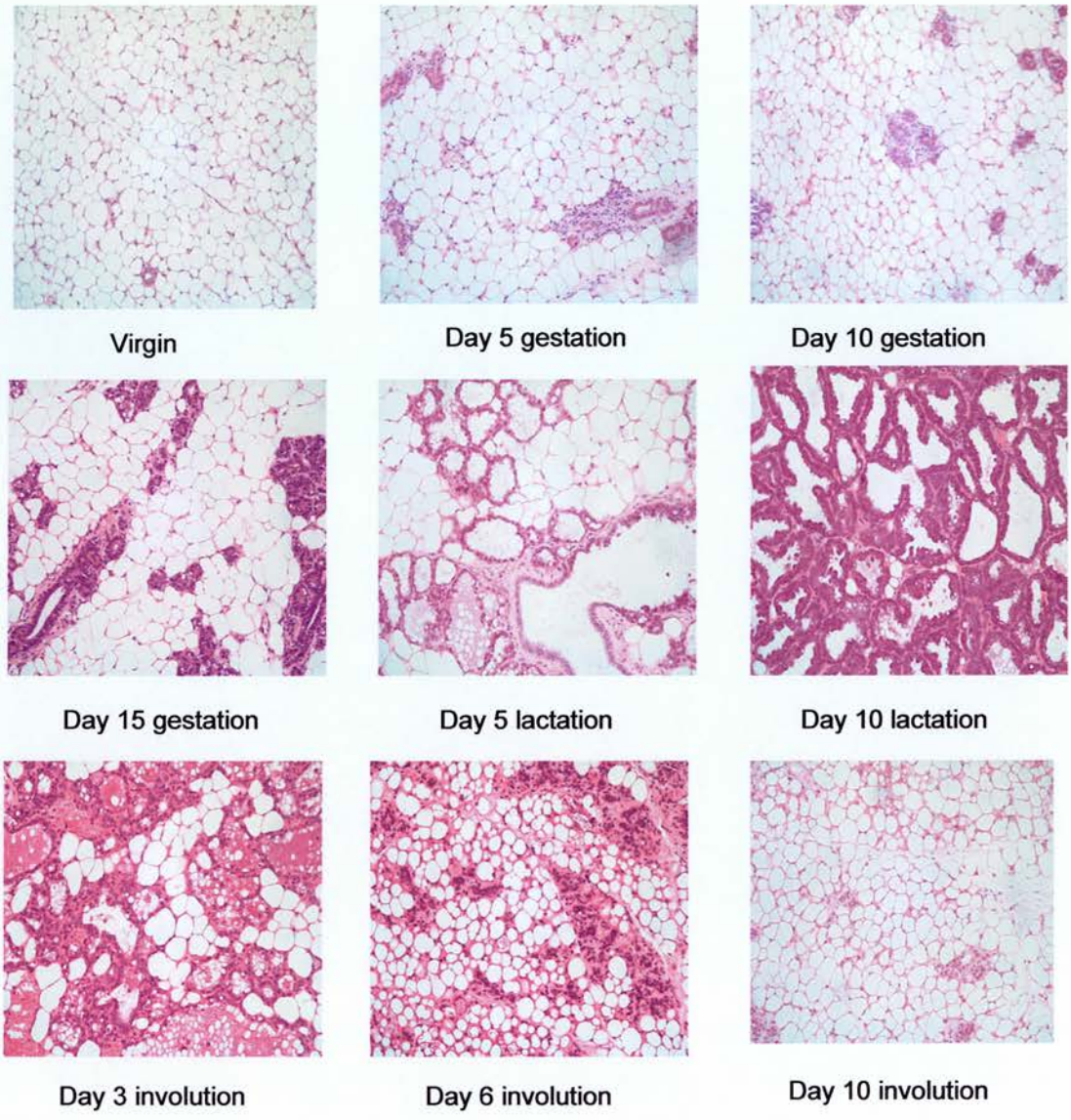


Figure 1.5

Mammary Gland Development (H&E staining)

In newborn mice a single primary duct is present until the onset of puberty when pronounced ductal growth occurs. Development of the ducts continues in cycling virgins leading to the formation of a ductal tree which fills the entire mammary fat pad. Extensive ductal branching and alveolar growth occurs during pregnancy. Terminal differentiation of the alveolar epithelium is completed at the end of gestation with the onset of milk secretion at parturition. After weaning the entire alveolar epithelium undergoes apoptosis and the gland is then remodeled. Within a few weeks the gland has the appearance of that of a mature virgin

animals. This characteristic has several ramifications; the tissue can be easily manipulated, and reasonable amounts of tissue are available for analysis. Furthermore, genetic manipulations whose consequences in other tissues would result in lethality can be studied because mammary cells transplanted into the cleared fat pad continue to grow and divide very effectively. Thus, normal virgin ductal or lobuloalveolar cells will readily take, grow and fill the fat pad with mammary ducts in virgin mice (or differentiate into alveolar cells if the recipient mice become pregnant). In virgin mice or rats, the implanted cells take about 8 weeks to fill the fat pad. The cleared fat pad also serves as a transplant site for preneoplastic and neoplastic mammary cells. The ability to transplant into the fat pad allows one to examine the morphogenic and tumourigenic capabilities of unknown cell populations. This is true for cells removed from the *in situ* mammary gland e.g. from the mammary gland of a late embryonic lethal knockout animal and for cells grown *in vitro*. Second, it is possible to generate chimeric glands composed of tissues from knockout and wild-type animals. Alternatively, epithelium and stroma can be separated enzymatically and recombined. The assembled tissues can also be grown as grafts under the kidney capsule of appropriate hosts (Cunha, 1994).

The ductal growth observed in mouse mammary glands is known to be closely controlled by ovarian steroid hormones (Daniel *et al*, 1987). Removal of the ovaries results in mammary gland regression, a phenomenon which can be reversed with the administration of oestrogen or progesterone (Mixner and Turner, 1942). Oestrogen receptor (ER) knockout mice have confirmed that this hormone is absolutely required for ductal outgrowth (Korach *et al*, 1996). Progesterone receptor (PR) null mice are infertile due to a failure to ovulate so the role of the PR has been assessed using ovariectomised mice. After hormone treatment control mice displayed extensive ductal branching but this was virtually absent in the PR-deficient mice showing that PR is also required for proper ductal growth of the gland. Prolactin is essential for the proliferation and functional differentiation of the lobulo-alveolar structures during pregnancy (Topper and Freeman, 1980). In the absence of the PR from the mammary epithelium, ductal side-branching fails to occur. This defect can be overcome by ectopic expression of the protooncogene *Wnt-1* and transplantation of

mammary epithelia from *Wnt-4-/-* mice shows that Wnt-4 also has an essential role in side-branching early in pregnancy (Brisken et al, 2000). *PR* and *Wnt-4* mRNAs colocalise to the luminal compartment of the ductal epithelium where progesterone induces Wnt-4 in mammary epithelial cells and is required for increased Wnt-4 expression during pregnancy. Thus, Wnt signalling is essential in mediating progesterone function during mammary gland morphogenesis. Prolactin signalling in the mammary gland also operates via the JAK/STAT pathway (Liu *et al*, 1997), predominantly through the two STAT5 isoforms, Stat5a and Stat5b (STAT5 was originally called mammary gland factor). Other factors including GH, epidermal growth factor and erythropoietin also use STAT proteins to regulate the transcription of specific genes through the JAK/STAT pathway (Schindler and Darnell, 1995). The STAT pathway is activated through ligand binding triggering dimerisation or oligomerisation of receptors. Receptor-associated tyrosine kinases (JAKs) cross-phosphorylate each other as well as the tyrosine residues on the receptors. Subsequently, SH2-containing latent cytoplasmic proteins from the STAT family are recruited to the receptor complex and phosphorylated by the JAKs. Two STAT proteins dimerise, translocate into the nucleus, and activate gene transcription by binding to IFN-γ activation sites in gene promoters (**Figure 1.6**).

The ability of individual receptors to activate overlapping but distinct sets of homo- and heterodimerising STAT proteins is thought to contribute to their signal specificity, in much the same way as the Smads. Animals deficient in STAT proteins are known to have a number of defects including impaired mammary gland development (Chapman et *al*, 1999, Akira, 1999, Udy *et al*, 1997). One of the most exciting finds in recent months is the observation that this complex cytokine-activated pathway can interact at the nuclear level with the Smad pathway (Nakashima *et al*, 1999) and this finding forms the basis for my investigation into Smad-STAT interactions within the mammary gland. Other important molecules involved in mammary gland development are colony-stimulating factor-1 (CSF-1) and oxytocin (OT). A natural mutant of the CSF-1 gene (observed as a consequence of osteopetrosis homozygosity in a mutant mouse) has, amongst other things, a severe lactational phenotype with a complete absence of milk secretion (Pollard and

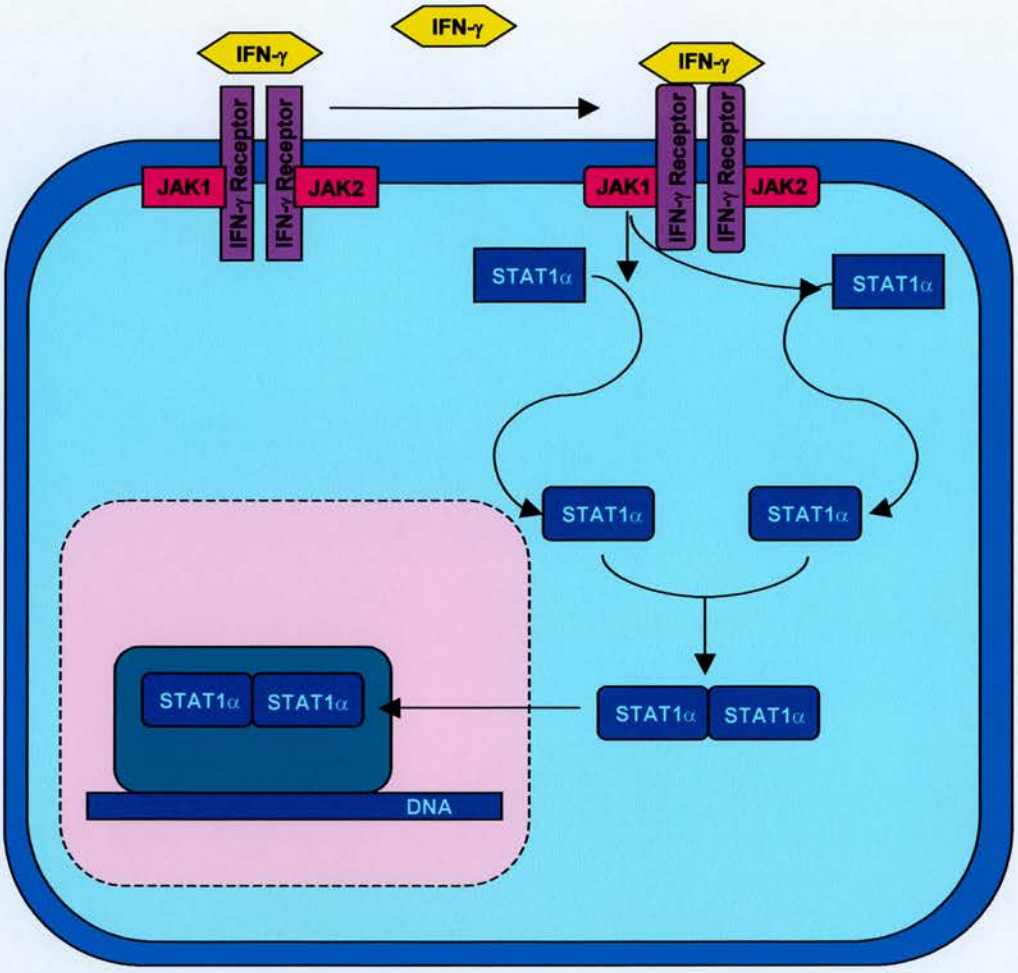


Figure 1.6

The Jak-STAT Signalling Pathway

The STAT pathway is activated through ligand binding triggering dimerisation or oligomerisation of receptors. Receptor-associated tyrosine kinases (JAKs) cross-phosphorylate each other as well as the tyrosine residues on the receptors. Subsequently, SH2-containing latent cytoplasmic proteins from the STAT family are recruited to the receptor complex and phosphorylated by the JAKs. Two STAT proteins dimerise, translocate into the nucleus, and activate gene transcription by binding to IFN-γ activation sites in gene promoters. **Figure 1.6** above utilises IFN-γ signalling via STAT1 homodimers as an example of this signalling pathway.

Heninghausen, 1994). OT knockout mice also fail to produce milk and this then leads to a rapid involution of the gland (Young *et al*, 1996, Nishimori *et al*, 1996 and Wagner *et al*, 1997). Because TGF-β is a potent inhibitor of epithelial cell proliferation TGF-β is assumed to be important during the involution stage of mammary gland growth and development. However, different TGF-β isoforms have been implicated in other areas of mammary gland biology including in regulating functional differentiation or the onset of milk secretion (Daniel and Robinson, 1992, Robinson *et al*, 1991). The role of TGF-β in involution of the mammary gland has been well characterised through a number of methods including *in vitro* work and knockout and transgenic animals. At the time the work presented in this thesis was initiated nothing was known about Smads in the mammary gland. However, recent papers have shed some light on the role that Smads play in the mammary gland, suggesting that these molecules are key players in this complex, dynamic tissue.

1.4.2 Smads and the Mammary Gland

The initial idea that Smads are involved in mammary gland biology derives from the fact that they have been elucidated as the downsteam signalling components of the TGF-β pathway, and because TGF-β has a clearly established role in mammary gland involution.

The evidence to date that Smads are involved in mammary gland biology comes from only a few published experiments showing that Smads are present in mammary carcinoma cell lines (Pouliot and Labrie, 1999) and that expression levels of the Smads become elevated and nuclear when mammary carcinomas are treated with perillyl alcohol, an anti-cancer agent currently in trials. These results demonstrate that the TGF-β signalling pathway was activated in the regressing carcinomas (Ariazi *et al*, 1999). Taken together, the characterised interaction between Smads and STATs – known to be important in mammary gland biology - and the clear role for TGF-β in extracellular matrix modelling suggests therefore that it seems entirely reasonable that the Smad family has an important role to play in normal mammary gland

function and development, and further investigation into this formed the basis of this research project.

1.5 Animal Models of Gene Function

Transgenic animal technology, and the use of germline manipulation for the creation of targeted gene mutations, has resulted in a plethora of murine models for biological research. Our understanding of some of the important issues regarding the mechanisms controlling cell division, differentiation and death has dramatically advanced in recent years through exploitation of these techniques to generate transgenic mice. In particular, the generation of mice with targeted mutations in genes encoding proteins of interest has proved to be a useful way of elucidating the function of these gene products *in vivo*. Transgenic mouse models have provided some insight into the complex events contributing to cellular dysregulation and the loss of growth control that can lead to tumorigenesis. They have also helped elucidate crucial roles for genes in development and normal tissue homeostasis.

The use of transgenic mice in biomedical research has now become widespread, and numerous applications have been developed. Several general reviews have been published, including those by Palmiter and Brinster (1986), Jaenisch (1988) and Hanahan (1989). Two of the most common uses of transgenic mice have been for (1) studies of tissue-specific and developmental stage-specific gene regulation and (2) for experiments of the phenotypic effects of transgene expression. Most cloned genes introduced into the mouse germ line have shown appropriate tissue-specific and stage-specific patterns of expression despite their integration into apparently random sites in the host genome. Gene transfer into the mouse embryo has therefore provided the definitive experimental assay to define the cis-acting DNA sequences that dictate specific patterns of transcription in the developing animal. In addition, the characterisation of regulatory sequences through such studies has provided tools for the second category of experiments in which the expression of gene products of interest is directed to specific sites in the developing mouse.

The first report of the direct introduction of new genetic material into the mouse embryo actually predates the widespread use of recombinant DNA techniques. In 1974, Rudolf Jaenisch and Beatrice Mintz found that when purified SV40 DNA was injected into the blastocoel cavity of mouse blastocysts, viral DNA sequences could be detected in somatic tissues of many of resulting animals, suggesting that they had integrated into the genome of embryonic cells. In addition, Jaenisch (1976) discovered that Maloney murine leukaemia virus could be stably introduced into the germ line by viral infection of pre-implantation mouse embryos. However, these studies did not immediately lead to attempts to introduce cloned eukaryotic genes into the germ line.

In 1980 it was reported that the microinjection of the cloned herpes simplex virus (HSV) *thymidine kinase* (*tk*) gene into the nuclei of cultured fibroblasts led to the stable incorporation and expression of the *tk* gene in 5-20% of the recipient cells (Anderson *et al*, 1980, Capecchi 1980). This finding suggested that the microinjection of DNA into the one cell mouse embryo might allow the efficient introduction of cloned genes into the developing mouse and led a number of investigators to test this possibility. Gordon *et al* reported the first successful introduction of a cloned gene into mouse somatic tissues by pronuclear injection in 1980. Shortly thereafter, several groups were successful in introducing cloned genes into somatic tissues as well as into the germ line by this technique (Brinster *et al*, 1981, Costantini and Lacy, 1981, Gordon and Ruddle, 1981, Harbers *et al*, 1981, and Wagner *et al*, 1981).

1.5.1 Gene Targeting in ES cells

Our understanding of the signal transduction network has been aided by the generation of mice that carry mutations in a cloned gene following homologous recombination in embryonic stem cells (ES cells). This is a more directed approach than the analysis of pre-existing mutants, as the molecule of interest has often already been analysed *in vitro*. A significant advantage of this approach is the identification of mutants that do not exhibit an overt phenotype, and which therefore

would not have been identified by classical genetics. Homologous recombination in ES cells is now a routine technique that is used to modify the mouse genome at any chosen locus. The principles of this technique were developed in the 1970s in yeast where contrary to the situation in mammalian cells, the majority of recombinations between introduced vector DNA and genomic DNA occur by homologous recombination as opposed to random integration.

Homologous recombination in mammalian cells between an artificial targeting vector and an endogenous gene was first achieved by Smithies *et al.* (1985) for the β-Globin locus, albeit at very low frequency. In 1981 two groups (Evans and Kaufmann, 1981 and Martin 1981) derived pluripotent embryonic stem cell lines from mouse blastocysts. Bradley *et al* (1984) were able to show that even after prolonged tissue culture such ES cells have the capacity to colonise the germ line of chimeric mice when injected into blastocysts. These experiments paved the way to altering the mouse genome by homologous recombination in ES cells. This was first achieved for the selectable hypoxanthine phosphoribosyl transferase (HPRT) gene locus (Doetschman *et al*, 1987, Thomas and Capecchi 1987) and subsequently targeting of non-selectable genes such as *int-2* and *c-abl* (Mansour *et al*, 1988, Schwartzberg *et al*, 1989) became possible after enrichment strategies for homologous recombination had been developed. The first germline transmission in mice of a targeted allele involved the corrected *Hprt* allele here in Edinburgh by Thompson *et al* in 1989.

ES cell lines are derived from pluripotent, uncommitted cells of the inner cell mass of pre-implantation blastocysts and can contribute efficiently to both somatic and germ line tissues after re-introduction into blastocysts. This capacity crucially depends on culture conditions that keep the ES cells in an uncommitted, undifferentiated state. These differentiation-inhibiting signals can be provided by feeder cells that also serve as a matrix for ES cell adherence and/or by the addition of leukaemia inhibitory factor (LIF) to the culture medium (Pease and Williams, 1990). The genetic background of the majority of available ES cell lines is 129, a strain from which ES cells lines are most readily established (Kawase *et al*, 1994). ES cells have been well characterised since their discovery in 1981 and a variety of protocols

exist to maintain their pluripotency throughout a gene targeting experiment. The key to good, undifferentiated ES cell culture is the addition of LIF to the culture medium – the first ES cells to be isolated were maintained on a feeder layer of mitomycin C-treated fibroblasts (Evans and Kaufmann, 1981 and Martin, 1981). However, the presence of feeders can complicate the selection processes for isolation of genetically modified cells (Hooper, 1987) and the discovery that medium conditioned by Buffalo rat liver cells (BRL) can substitute for feeders greatly improved experiments (Smith and Hooper, 1987). It has since been elucidated that an isolated, purified factor called LIF (leukaemia-inhibiting factor) will also effectively maintain ES cell pluripotency (Williams *et al*, 1988 and Smith *et al*, 1988).

The first step in a gene targeting experiment is to isolate a genomic clone or clones containing the gene of interest. The targeting vector should be constructed from a genomic clone of the same genetic background as the ES cells to be used, known as isogenic DNA. Once the genomic clone has been isolated and characterised a targeting vector must be built, usually one of two forms: a replacement vector or an insertion vector. The replacement vector is the most commonly used type of approach and is linearised in such a way that the vector sequences remain collinear with the target sequences. Chromosomal sequences are replaced by vector sequences by a double crossover event involving the flanking homologous region. The success of this homologous recombination is dependent on several parameters; the total length of the homologous region (0.5kb minimum, 10kb maximum for ease of handling) (Hasty *et al*, 1991, Deng and Capecchi 1992), recombination rates can increase with the use of isogenic DNA (te Riele *et al*, 1992) and absolute frequencies of homologous recombination seem to be locus dependent. Normally a selection marker is incorporated into the targeting vector in order to screen for successfully targeted clones – the most commonly used marker is a cassette carrying the neomycin resistance gene (neo^r) under the control of a strong promoter. However, to enrich for clones that have undergone homologous recombination as opposed to random integration a strategy termed "positive/negative" selection was developed by Capecchi and co-workers (Mansour *et al*, 1988). A thymidine kinase (TK) gene from HSV is inserted at the end of the linearised targeting construct. Cells that have

undergone homologous recombination will have lost the TK gene, whereas cells in which the construct integrated randomly can be eliminated using a toxic nucleoside analogue such as ganciclovir. Enrichments using this technique are typically from 3 to 10 fold. Screening for correctly targeted clones is by either Southern Blot or PCR with one primer derived from the newly introduced selection cassette and the other primer hybridising to genomic sequences outside the targeting construct.

Although the generation of a simple knockout as described above can be a relatively simple process it can also be one strewn with problems and pitfalls. Incomplete knockouts can be created when coding sequences remain present in the genome and form truncated or mutant forms of the protein. This can be caused by read-through transcription through the inserted selection marker followed by initiation of translation from downstream AUGs (for example see Tanaka *et al*, 1997) or splicing which may lead to the expression of modified proteins (for example see Li *et al*, 1992). Removal of all coding regions or regulatory elements of genes can also cause problems as it may lead to large genomic deletions resulting in the loss of regulatory elements governing the expression of unrelated genes – this can be problematical especially with clustered genes. Here different approaches to the targeting experiment with the same gene can lead to a vast range in phenotypes dependant on what regulatory factors were deleted at the same time (for example see Olson *et al*, 1996). Selection cassette interference may also cause problems as these markers of recombination are almost always driven by strong promoters which may disrupt the expression of neighbouring genes (for example see Fiering *et al*, 1995). Of course one of the other major drawbacks with this technology can be that the phenotype of the knockout is so severe that little biology can be interpreted other than the extreme importance of the targeted gene at some point during development. However, many genes are switched off after embryonic development or are limited to tissue specific expression and so analysis by this means often proves fruitless. Recently strategies exploiting site-specific DNA recombination have been incorporated into transgenic and gene targeting procedures to allow *in vivo* manipulation of DNA in ES cells or living animals. In other words it is possible to generate a deletion of the gene of interest at a specific point in time and in a specific tissue.

1.5.2 Cre-Lox Technology

Cre is the 38kDa product of the *cre* gene of bacteriophage P1 (Sternberg, 1978) and is a site-specific DNA recombinase. Within the bacteriophage Cre has two roles, firstly to circularise the viral genome should this fail to be carried out by host cell recombination machinery (Segev and Cohen, 1981) and secondly to maintain correct unit copy segregation of the prophage by resolving any dimeric molecules into monomers prior to cell division (Austin *et al*, 1981). Cre recognises a 34 base pair site on the P1 genome called *loxP* (locus of X-over of P1) and efficiently catalyses reciprocal conservative DNA recombination between pairs of *loxP* sites (Austin *et al*, 1981). The loxP site consists of two 13 base pair inverted repeats flanking an 8 base pair non palindromic core region (Hoess *et al*, 1982) that gives the loxP site an overall directionality that, by convention, is depicted as an arrow. Cre mediated recombination between two directly repeated loxP sites results in excision of the DNA between them as a covalently closed circle (**Figure 1.7**). Unlike many other recombinases no accessory factors or DNA topological requirements are required for efficient Cre-mediated DNA recombination and this makes this system uniquely suited to genomic manipulation of eukaryotic cells (Sauer, 1988). The strategy is similar in many ways to that described earlier for conventional gene targeting – the target gene is modified by homologous recombination in ES cells so that it is flanked by two *loxP* sites. The correctly targeted ES cells are injected into blastocysts before being introduced into pseudopregnant mice and animals containing the modified gene (which are at this point phenotypically wild type) are then crossed with mice expressing Cre in the desired target tissue and Cre-mediated excision results in a tissue-specific gene ablation. This strategy was first demonstrated by Gu *et al* (1994) using a mouse in which the promoter and first exon of the DNA Polymerase β gene (*polβ*) were flanked by *loxP* sites. When mated to a transgenic mouse that specifically expressed Cre in T cells, the *polβ* gene was inactivated in the T cell population, but not in any other tissue.

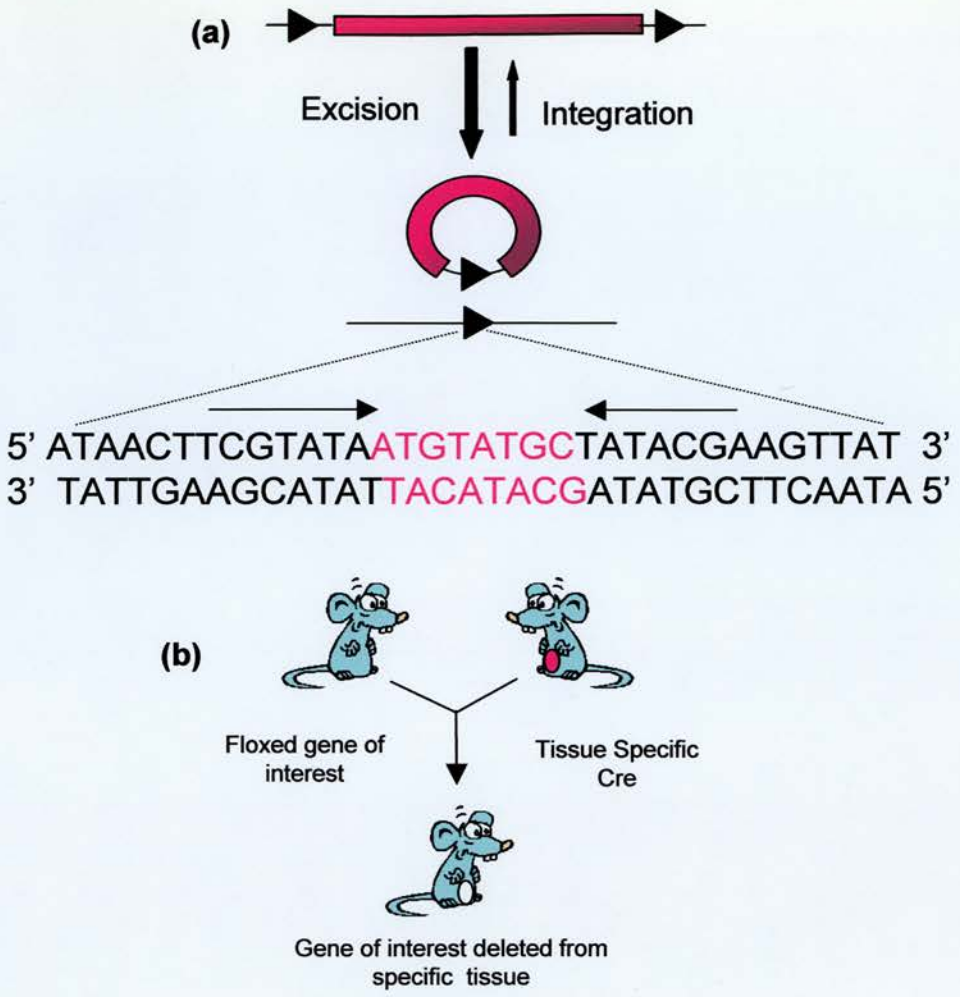


Figure 1.7

Conditional Gene Targeting

(**a**) Cre recognises a 34 base pair site on the P1 genome called *loxP* (locus of X-over of P1) and efficiently catalyses reciprocal conservative DNA recombination between pairs of *loxP* sites. The loxP site consists of two 13 base pair inverted repeats flanking an 8 base pair non palindromic core region that gives the loxP site an overall directionality that, by convention, is depicted as an arrow. Cre mediated recombination between two directly repeated loxP sites results in excision of the DNA between them as a covalently closed circle (**b**) Correctly targeted ES cells are injected into mice and animals containing the modified gene (which are at this point phenotypically wild type) are then crossed with mice expressing Cre in the desired target tissue and Cre-mediated excision results in a tissue-specific gene ablation

A particularly powerful feature of a conditional gene inactivation strategy using Cre is that the same *loxP*-tagged mouse can be used for gene ablation independently in a large number of different tissues, or at different developmental times, by simply mating it with a corresponding Cre transgenic that displays the desired tissue or temporal specificity of expression. Thus the same genetically modified animal can be used to answer a variety of different questions relating to the expression and function of the target gene. Of course, crucial to the success of this strategy is the prior careful evaluation of the pattern of expression of the *cre* transgenic animal. Mosaic expression or exogenous tissue expression would defeat the purpose of the experiment and so careful choice of promoter must be considered. There are currently many Cre transgenic mice available ranging from oocyte-specific (Lewandoski *et al*, 1997), mammary gland (Wagner *et al*, 1997 and Selbert *et al*, 1998), liver (Kellendonk *et al*, 2000) and pancreas (Ray *et al*, 1999). This method of introducing Cre is tissue specific but relies very heavily on the availability of suitable tissue-specific promoters and not all attempts at achieving tissue-specific expression of Cre recombinase are successful; for example a transgenic mouse strain engineered to express Cre under the control of the platelet endothelial cell adhesion molecule-1 promoter (PECAM-1) was predicted to express Cre only in endothelial cells (Terry *et al*, 1997). However, when the PECAM-1/Cre mouse was crossed to a floxed mouse, the progeny were seen to have undergone recombination in all tissues. The conclusion from this work is that if the promoter is even briefly active during an early stage of development this can be sufficient to induce floxing in a non-tissue-specific manner. An attractive alternative to native promoters is to use synthetic inducible systems to control *cre* expression. The elegant tetracycline-regulated transcriptional systems (Gossen *et al*, 1992 and 1995) offer the possibility of inducing *cre* expression at a desired time either by simply dosing an animal with tetracycline or by withdrawing animals from tetracycline administration. A much more rapid approach to the problem can be possible by viral-mediated gene transfer of Cre recombinase (Anton and Graham, 1995). This uses a basic strategy where sections of the wild type virus genome, such as sequences encoding the E1 and E3 genes are replaced with a Cre expression construct. Such recombinant viruses are replication deficient but still capable of host cell infection. Many groups have used

this approach successfully. As cultured cells can be exposed directly to the virus very high efficiencies of infections can be achieved, more so than with standard transfection efficiencies. For example infection of CV1 cells with adenovirus Cre (AdCre) resulted in recombination mediated gene activation in up to 100% of the cells (Kanegae *et al*, 1995). The technology has also been used successfully *in vivo* where the commonest way of administering the AdCre is by intravenous injection (IV), although this route does not lead to uniform infection of somatic tissues. Cre-expressing adenovirus vectors can be administered by IV to loxP-modified mice and cause high levels of recombination in the primary targets tissues of liver and spleen (Wang *et al*, 1996 and Rohlmann *et al*, 1996).

The TGF-β superfamily is one of the most studied groups of proteins by gene inactivation in animals. Knockout mice have been generated for TGF-β1, β3 and some of their receptors. Null mice have also been characterised for many of the activins and their receptors and also the BMPs and their receptors (for review see the Biomednet mouse knockout database at http://www.biomednet.com/db/mkmd. Following the discovery of the Smads in 1996 there followed a spate of Smad knockout mice including the embryonic lethal Smad4 (Sirard et *al*, 1998), embryonic lethal Smad2 (Weinstein *et al*, 1998), colorectal cancer prone Smad3 (Zhu *et al*, 1998) and the embryonic lethal Smad5 (Chang *et al*, 1999). The high incidence of embryonic lethality associated with Smad null mice suggests that they are very important in the early stages of development. However, this fact has led to much interest in producing conditional deletions of the Smads in order to learn more about their tissue-specific roles in the mammal.

1.6 Aims of the Project

The broad aim of this project is to further understand the role that Smad4 plays in normal mammary gland function and development. In order to achieve this goal several approaches have been taken; the first part of the project examines how Smad overexpression affects embryonic stem cells. These cells are known to not express

various receptors for the TGF-β family until they become differentiated so they represent a method of examining how overexpression of non-activated downstream signalling components can affect the whole pathway.

The second approach has been to generate a mammary gland-specific Smad4 transgenic mouse. This mouse contains a stable transgene consisting of the BLG promoter upstream of a FLAG epitope-tagged human Smad4 cDNA. This transgenic should be useful in again addressing the question of what role Smad4 plays in the lactation to involution progression and to what extent it is involved in the apoptosis of the gland.

The third and possibly the most exciting has been the production of a "floxed" Smad4 mouse. The embryonic lethality associated with conventional gene targeting of the Smad4 locus has prompted the generation of a conditional targeting strategy whereby Exon I of the Smad4 locus can be deleted in a tissue-specific manner in the presence of Cre recombinase. This "floxed" mouse will be crossed to a BLG-mammary gland specific Cre transgenic (Selbert *et al*, 1998) which has been characterised to express Cre protein only in the epithelial cells of the gland during pregnancy and lactation. This model will allow us to answer the question of what role Smad4 plays in the massive epithelial cell apoptosis characteristic of involution of the mammary gland.

The final approach has been an examination of Smad interactions with other known signalling pathways in the gland. Here I have examined first the normal expression pattern of Smads in the mammary gland and then gone on to characterise their *in vivo* interactions with members of the STAT pathway, a group of proteins known to be important in the mammary gland and known to interact with Smads *in vitro*.

Chapter 2 - Materials and Methods

2.1 Preparation of DNA

2.1.1 Large-scale preparation of plasmid DNA – Maxiprep

Two 50µl aliquots of competent *Escherichia coli* cells (>2x10^7 CFUµg pUC19) of the DH5α (Life Technologies) were removed from -70°C storage and thawed on ice. 1µl (approximately 1µg) of DNA solution was added to one of the tubes of *E. coli* cells and the tubes incubated on ice for 30 minutes. The cells were heat shocked at 42°C for 45 seconds and then placed on ice for two minutes. 800µl of LB medium was added to the tube and the cells were incubated at 37°C for 1 hour in a shaking incubator (Gallenkamp) (220 rpm). The transformed bacteria were plated out onto L-amp plates (disposable 100mm bacterial culture dishes containing L-broth (See **Appendix A**) supplemented with 50µg/ml ampicillin (Sigma) and 1.2% Bactoagar and incubated overnight at 37°C.

The following day a colony was picked from the plate transformed with plasmid DNA (the other control plate should produce no colonies) and was placed in 10mls L-broth supplemented with ampicillin and the cells grown for about 8 hours in a shaking incubator at 37°C. This culture was added to a 500ml flask of L-broth supplemented with ampicillin and incubated overnight in a shaking incubator at 37°C, 220rpm.

The following morning the cells were harvested by centrifugation at 6000xg for 10 minutes at 4°C. The supernatant was discarded and the cells lysed by resuspending in 10mls Solution P1 of the Qiagen maxiprep kit (Qiagen) (See **Appendix A**).

10 mls of solution P2 (See **Appendix A**) was then added, the lysate mixed gently but thoroughly by inversion and left to incubate at room temperature for 5 minutes

10 mls of chilled Solution P3 (See **Appendix A**) was added and the lysate mixed gently by inversion before incubation on ice for 20 minutes

The precipitated proteins were removed by centrifugation at 20,000xg for 30 minutes at 4°C and the supernatant containing the plasmid DNA removed promptly. A Qiagen-tip was equilibrated with Buffer QBT (See **Appendix A**) and the plasmid-containing supernatant was allowed to enter the column resin by gravity flow. The Qiagen-tip was then washed twice with Buffer QC and then the plasmid DNA was eluted with 15mls Buffer QF (See **Appendix A**).

The plasmid DNA was precipitated by adding 0.7 volumes of isopropanol, mixing by inversion and incubating for 10 minutes at room temperature. The DNA precipitate was recovered by centrifugation at 17000xg for 30 minutes at 4°C. The supernatant was discarded and the pellet was washed with 5mls of 70% ethanol at room temperature and then dried under vacuum. The plasmid DNA was resuspended in 200µl ddH_2O and quantified by spectrophotometry (Genequant, USA). Plasmid DNA was stored at -20°C.

2.1.2 Small-scale preparation of plasmid DNA – Miniprep

Isolation of plasmid DNA from bacterial cultures for the verification of plasmid identity by restriction digestion was carried out using the QIAprep Spin Plasmid Kit (Qiagen). Single colonies were picked from L-amp plates and inoculated into 5mls LB medium containing the appropriate antibiotics (100µg/ml ampicillin) in Falcon 2059 tubes (Becton Dickinson). After an incubation period of 14-16 hours at 37°C in an orbital shaker at 220rpm, a final 1.5 ml sample from each of the resulting late logarithmic phase cultures was transferred to a microfuge tube and the bacteria pelleted by microcentrifugation at 6500rpm for 5 minutes. The supernatant was

discarded and the bacterial pellets resuspended in 200µl Cell Resuspension Buffer (see **Appendix A**), lysed by the addition of 200µl of Cell Lysis Buffer (see **Appendix A**). This solution contains alkaline sodium dodecyl sulphate (SDS) and this denatures bacterial proteins and the strong alkaline conditions used in this procedure cause denaturation of both plasmid and chromosomal DNA. This was mixed gently by inversion and then neutralised by addition of 200µl Neutralisation solution (see **Appendix A**). This neutralisation caused rapid renaturation of circular plasmid DNA whereas most chromosomal DNA, bacterial protein and SDS (as potassium dodecyl sulphate) formed a white gelatinous precipitate which was separated from the aqueous, plasmid-containing supernatant upon microcentrifugation at 13000 rpm for 5 minutes. This supernatant (cleared lysate) was promptly separated from the debris by pipetting into a clean Eppendorf microfuge tube.

1-2µl samples of DNA purified from cleared lysates was digested with appropriate restriction enzymes and analysed by agarose gel electrophoresis.

2.2 Manipulation of DNA

2.2.1 Restriction Endonuclease Digestion of Plasmid DNA

Plasmid DNA was digested in sterile Eppendorf microfuge tubes containing a mixture of DDW, the appropriate restriction endonuclease and the recommended correct ionic strength buffer. Boehringer or New England Biolabs supplied all restriction enzymes. 10x stock buffer solutions were used diluted 1:10 in the final reaction mix. Restriction enzymes were stored at –20°C in buffers containing 50% glycerol and kept on ice while in use. Sufficient enzyme was used to ensure complete digestion. 1 unit of enzyme being the amount required to digest 1µg of pBR322 plasmid DNA to completion in 1 hour. The amount of enzyme was kept below 10% of the final reaction volume as glycerol can interfere with enzymatic activity. Digestion was carried out at 37°C for most enzymes for a minimum of one hour.

2.2.2 Agarose Gel Electrophoresis of DNA

Agarose gel electrophoresis was carried out in order to confirm digestion, characterise plasmids or purify restriction fragments. DNA molecules are separated by electrophoresis according to size, conformation (supercoiled, nicked circular or linear), agarose concentration and applied current. Tris-borate EDTA (TBE) was routinely used as an electrophoresis buffer (See **Appendix A**).

Preparation of gels and electrophoresis in BioRad MiniSub cells was carried out as described by Sambrook *et al.* (1989). Samples were loaded by adding 0.1 volume gel loading buffer and applied to the wells of a gel submerged in electrophoresis buffer. Fragment sizes were estimated with reference to a 1-kilobase ladder (Life Technologies) See **Appendix A**.

2.2.3 Phenol: Chloroform Extraction and Ethanol Precipitation of DNA

These procedures were carried out as described in Sambrook *et al.* (1989) in order to remove proteins and to concentrate DNA. The resulting DNA pellets were then briefly air-dried and redisolved in an appropriate volume of double-distilled water.

2.2.4 Estimation of DNA Concentration and Purity

Two methods for the estimation of the concentration of DNA solutions were used: a spectrophotomeric method and an ethidium bromide fluorescent quantitation method (Sambrook et al, 1989). DNA samples from large and small-scale DNA preparations were quantitated by ultraviolet spectrophotometry.

The optical densities (OD) of 1:100-diluted DNA solutions were measured in a Genequant II RNA/DNA Calculator (Pharmacia) at λ= 260nm. An OD_{260} of 1 was taken to be equivalent to a concentration of 50μg/ml of double-stranded DNA in the

diluted sample. The ratio of the ODs at λ=260nm and at λ=280nm is a measure of the purity of a DNA sample. Accordingly, $OD_{260/280}$ ratios measured in the Genequant were recorded for DNA samples.

The concentration of certain DNA solutions was estimated by ethidium bromide fluorescence method. Small samples of DNA (e.g. 0.5µl) and a range of serially diluted kilobase ladder were spotted onto a plastic Petri dish containing 1% agarose gel and 5µg/ml ethidium bromide. After 10-15 minutes to allow absorption of DNA onto the gel, the plate was exposed on an ultraviolet transilluminator and the DNA concentration was estimated by comparison of the fluorescent intensity of the sample against that of the standards.

2.2.5 Isolation of DNA fragments from agarose gels

Digested DNA fragments were visualised by UV transillumination and then excised from 0.7% agarose TBE buffered gels and purified using the QIAEX Gel extraction kit (QIAGEN) according to manufacturers instructions.

2.2.6 Modification of the ends of linear DNA: dephosphorylation

1 unit shrimp alkaline phosphatase (SAP) (United States Biochemicals USB) was added to 40µl gel-purified DNA together with the appropriate dephosphorylation buffer and distilled water up to a final volume of 50µl. This reaction was incubated at 37°C for one hour and then the phosphatase was inactivated by heating to 65°C for 15 minutes.

2.2.7 Klenow reaction

The large fragment of DNA polymerase I, Klenow enzyme (United States Biochemical) retains 5' to 3' polymerase activity but lacks any exonuclease activity. This polymerase was used to blunt 5' overhangs of DNA generated by restriction

enzyme cleavage when needed. The enzyme (1-5 units, depending on reaction volume) and 0.1 volumes of 0.5mM dNTP mix can be added directly to a restriction digest and incubated at 37°C for 30 minutes. If the blunted DNA was to be used in subsequent ligations the correct size band was excised from the gel, purified and SAP treated if necessary.

2.2.8 T4 DNA Polymerase

T4 DNA Polymerase catalyses the synthesis of DNA in the 5' to 3' direction so, like Klenow fragment, can be used to fill-in 5' overhangs. However, T4 DNA polymerase also has a 3' to 5' exonuclease activity that will remove any 3' overhangs generated by certain restriction enzymes to leave blunt ends. As with Klenow, the enzyme and other reaction components were added directly to restriction digests after they had gone to completion. Each reaction contained 3 units of polymerase (New England Biolab)/µg of DNA, 0.1 volumes of 1mM each dNTP mix (Pharmacia Biotech), 0.1 volumes of 0.5mg/ml BSA and 0.1 volumes of T4 reaction buffer

2.2.9 Southern Blot Hybridisation

10µg of genomic DNA was digested with the appropriate restriction enzyme. The sample was then run on a 0.7% agarose gel in 1 x TBE. The gel was then prepared for blotting onto a nylon membrane. If the fragments of interest were predicted to be larger than 10kb, transfer of the DNA to the nylon membrane was improved by partially depurinating the DNA prior to transfer. This was achieved by soaking the gel for 10 minutes in 0.2M HCl, followed by a rinse in ddw.

The DNA was transferred from the gel onto a positively charged nylon membrane, Zeta-Probe® GT (Bio-Rad) according to the manufacturers instructions using 0.4M NaOH as the transfer buffer. After blotting overnight the membrane was rinsed in 2 x SSC and air dried. The nylon membrane was then baked at 80°C for 30 minutes and stored between 3MM Whatman paper at room temperature until needed.

UltraHYB™ (Ambion) hybridisation buffer was heated to 68°C to ensure complete solubilisation of all components. 10mls was then removed and added into a hybridisation tube containing the baked nylon membrane. The membrane and UltraHYB™ buffer were then incubated at 42°C for 1 hour in a hybridisation oven (Hybaid)

DNA probe fragments were gel purified and 20ng labelled with 50μCi, 3000Ci/mol [α-^{32}P]-dCTP (ICN) using High Prime random oligonucleotide premixed solution. (Boehringer Mannheim). Unincorporated nucleotides were removed by passing labelled probes down a Sepahdex G50 column (NICK column – Pharmacia).

Labelled DNA probe was denatured by boiling and then immediately diluted in hybridisation solution overnight at 42°C. After overnight incubation the hybridisation solution was discarded and the filters washed in twice in 2xSSC, 0/1% SDS and twice in 0.1%SSC, 0.1% SDS. The filters were sealed in heat-sealable bags and exposed to BIOMAX MS-1 film in an autoradiography cassette with intensifying screens.

2.2.10 Dideoxynucleotide Chain Termination Sequencing of Plasmid DNA

Double stranded plasmid DNA template was prepared by the Qiagen maxi-prep method as described and then alkaline denatured. For each sample, 5μg DNA was diluted in 45μl DDW and then denatured with 5μl 2M NaOH, 2mM EDTA and incubated at ambient temperature for 5 minutes. The solution was neutralised with 2M ammonium acetate (pH 4.6) and the DNA precipitated at –70°C for 30 minutes with 185μl absolute ethanol. The DNA was recovered by microcentrifuge at 13000 rpm at 4°C for 10 minutes and dried in a vacuum dryer for 5 minutes. The DNA was resuspended in 6μl DDW.

The resuspended DNA was annealed to 2pmol primer in 1X Sequenase Reaction Buffer (United States Biochemicals) in a final volume of 10µl by heating to 65°C for 2 minutes and then allowed to cool to 25°C slowly in a waterbath. The annealed template/primer mix was then labelled at room temperature for 10 minutes with 5µCi (1000Ci/mmol) α-^{35}S]-dATP (ICN) and terminated at 37°C for at least 5 minutes with each of the four dideoxyribonucleotides using dGTP reagents and protocols provided with the Sequenase Version 2.0 DNA Sequencing Kit (United States Biochemicals). Each termination reaction was stopped using the provided "Stop" solution, which also acts as a dye-containing loading buffer.

Processed samples were denatured by heating to 75°C immediately before loading 3µl into wells created by sharks tooth combs at the top of a pre-run 6% denaturing polyacrylamide sequencing gels (See **Appendix A**). The gel was run using S2 sequencing apparatus (Life Technologies) and a model 3000 microcomputer electrophoresis power supply (Life Technologies) at 70W in 1X TBE.

2.3 Targeting Construct Preparation

2.3.1 Screening the λ PS mouse genomic library

A mouse genomic λPS library (MoBiTec) was screened as follows: A single colony of the bacterial strain C600 (-Cre) was grown in 50mls of LB++ in a 250ml flask at 37°C overnight. Next day the bacteria were transferred to a sterile 50ml tube and centrifuged to pellet the bacteria. They were then resuspended in 25mls of $MgSO_4$ and stored at 4°C.

The viral library was prepared as serial dilutions and 300µl of the C600 bacteria was infected with 1µl of various dilutions and incubated for 30 minutes at 37°C. Plates were prepared with LB/MgSO4 and LB/top agar was added to infected bacteria and poured onto the plates. These were left overnight at 37°C. The following day the plaques were counted to calculate the exact titer of pfus of the library stock.

The library was plated at $2x10^6$ pfu total and filter hybridisation was carried out according to standard protocols. Positive plaques were purified by additional rounds (usually 3) of plating and hybridisation. Single plaques were picked using a glass pipette and then transferred into a reaction tube containing 500µl SM/20µl chloroform.

The next step is in preparation of the Cre containing BNN132 cells (+Cre). From BNN132 cells grown on kanamycin plates (25µg/ml) a single colony was transferred into LB++ medium and prepared plating bacteria as described for the C600 bacteria. 20µl of the purified phage stored in SM/chloroform was added to 200µl of the BNN132 plating bacteria. This generates automatic sub cloning from the addition of Cre. This was incubated at 37°C for 30 minutes and then plated onto ampicillin plates and incubated overnight. The next day a single colony was picked and grown to saturation for maxiprep using standard protocols.

The result of screening the library was a 15kb genomic clone which encompassed exons1, 2 and 3 and upstream sequence from exon 1of the mouse Smad4 locus.

LB/$MgSO_4$ Prepare LB agar plates as described before. After autoclaving, add sterile 1M $MgSO_4$ to a final concentration of 10mM. Pour plates and store at 4°C.

LB++ broth Prepare LB broth as described before. Add sterile 1M $MgSO_4$ (1:100/v: v) to a final concentration of 10mM and 20% maltose (1:100/v: v) to a final concentration of 0.2%. Store at room temperature.

SM (per litre)

5.8g	NaCl
2.0g	$MgSO_4.7H_2O$
50ml	1M Tris-HCl, pH 7.5

Autoclave, add 5mls of 2% gelatin solution
Store at room temperature.

LB/top agar Prepare LB broth as described before. Add $MgSO_4$ (2.4g/L) to a final concentration of 10mM and agarose (7.2g/L) and autoclave. Store at room tempretaure.

2.3.2 Creating the floxed Smad4-HPRT targeting vector

The floxed Neomycin/TK cassette has been described before (Potocnik *et al*, 2000) and briefly consists of a neomycin antibiotic cassette and a thymidine kinase cassette with single loxP sites on either side of the cassette in the same orientation. The floxed Neo/TK cassette can be excised from its pBluescript backbone by restriction digest with BamHI and XhoI.

The single loxP sites used in the creation of the floxed Smad4 targeting construct were designed with EcoRV and BamHI sites on either end. These sites aid in cloning the loxP sites into the targeting vector and also help determine orientation of the loxP site after cloning. Other than these adaptations the loxP sites used are standard and do not contain any altered bases. The sequence of the loxP is:

GGATCCATAACTTCGTATAGCATACATTATACGAAGTTATGATATC

Dr Stefan Selbert at the Department of Pathology, Teviot Place, Edinburgh, created the floxed Smad4 targeting construct. The strategy employed was to isolate an 8kb BHI fragment from the original 15kb λ clone which was determined by restriction digest, southern blotting and sequencing to contain Exons I, II and III and some upstream sequence of the murine Smad4 gene. A single loxP site, as described above, was inserted into a unique NcoI site approximately 700bp upstream of Exon I. The floxed Neo/TK targeting cassette was inserted into a unique Bln I site approximately 1kb downstream of Exon I. The final step in generating the targeting construct was the introduction of an HPRT cassette into the NotI site of pBluescript

i.e. outside the 3' region of homology. The HPRT cassette was excised from pPOLYIII-Hprt using a Not I restriction digest. The pPOLYIII-Hprt cassette was a kind gift of Professor David Melton, Edinburgh University. The pTarVec plasmid was transfected into bacteria and grown to saturation so that maxipreps of the plamid could be obtained. 150μg of ClaI linearised plasmid DNA was used for electroporation into ES cells.

2.3.3 Constructing the BLG-Smad4 transgene

The human Smad4 cDNA used in the BLG-Smad4 transgene was a kind gift of Dr Rik Derynk, UCSF, USA. The cloning strategy employed to create the BLG-Smad4 vector was used by Dr Stefan Selbert, Medical School, University of Lubeck, Germany. The human Smad4 cDNA has been described previously (Zhang Y *et al*, 1996). The flag-tagged human Smad4 cDNA was cut out of the full vector with the restriction endonucleases ClaI and Hpa I (New England Biolabs). The resulting linear cDNA fragment was purified from a gel by the Quiaex Gel Extraction (QIAGEN) kit as described previously. The BLG-pBS/SK+ plasmid has been described previously (Selbert *et al*, 1998) and consists of the ovine beta lactoglobulin (BLG) promoter sequence cloned into pBluescript. This vector was opened with digestion with EcoRV (New England Biolabs) and the human Smad4 cDNA ligated downstream of the BLG promoter using Ready-to-Go T4 Ligase (Amersham).

2.3.4 Preparation of Transgene for Pronuclear injection

The BLG-Smad4 transgene was removed from the vector backbone by restriction digest with XhoI and NotI (New England Biolabs) and then purified by phenol chloroform extraction and ethanol precipitation as described. An equal volume of phenol: chloroform: isoamyl alcohol (25:24:1) was added to the sample and the solution was mixed well then centrifuged at 13,000rpm for 1 minute. The aqueous phase (upper layer) was pipetted into a clean microfuge tube, whilst the cloudy interface, containing the proteins and the lower organic phase were discarded. An

equal volume of chloroform was then added and the sample spun at 13,000rpm for 1 minute. The upper phase, containing the DNA was pipetted into a clean microfuge tube and the organic lower phase discarded. The purified DNA was then concentrated and sterilised by ethanol precipitation. For ethanol precipitation 2 volumes of ethanol and $1/10^{th}$ the volume of 3M sodium acetate (pH 5) were added to the DNA solution. The sample was placed at -20^0C for at least 1 hour and then centrifuged at 13,000rpm in a microfuge for 20 minutes. Ethanol was poured off and the pellet washed with 1ml of 70% ethanol. The sample was centrifuged for 15 minutes and the wash step repeated. The transgene DNA was resuspended in sterile injection buffer (10mM Tris, 0.1mM EDTA) at a concentration of 10μg/ml. The resuspended transgene was then dialysed against a large volume of injection buffer for 4 days at 4^0C, the buffer being changed daily. Following dialysis the transgene DNA was stored frozen in 20μl aliquots.

2.3.5 Collection of Fertilised Oocytes and pronuclear injection

This work was carried out by both Dr. Alastair Mackenzie in the MRC Transgenic Facility, Edinburgh and Dr Roberta James at the CRC labs, Western General Hospital, Edinburgh.

Fertilised oocytes were harvested from superovulated mice by standard methodology and placed in M2 medium containing hyaluronidase solution (300mg/ml) to remove the cumulus cells surrounding the fertilised oocytes, after a few minutes the oocytes were transferred to a petri dish containing fresh M2 medium and cultured at 37^0C prior to injection.

The BLG-Smad4 construct was prepared and injected into the pronuclei of the fertilised oocytes using a micromanipulator. Injected oocytes were then cultured overnight at 37^0C and the following day oocytes at the two-cell stage were selected and transferred into the oviducts of 0.5-day postcoitum pseudopregnant mice. Between 10-18 oocytes were transferred to each oviduct. For an in depth description

and discussion on this technique see "Manipulating the Mouse Embryo: A Laboratory Manual", 2nd Edition 1994 by Hogan, Beddington, Costantini and Lacy."

2.3.6 Extraction of DNA from Mouse Tails

Tail samples were taken from mice of 3 weeks of age and DNA extracted using the following method. Tails were incubated shaking overnight in 500µl of lysis buffer (See Appendix A) with 30µl of Proteinase K (20mg/ml).

The DNA was purified by phenol extraction. 500µl of TE saturated phenol was added and the tube shaken vigorously. The sample was then microfuged for 5 minutes and the aqueous phase carefully removed to a clean microfuge tube, avoiding the lower phenol phase and material at the interface. This procedure was followed again for 500µl of phenol: chloroform (1:1) and then 500µl of chloroform. DNA was precipitated by the addition of $1/10^{th}$ volume 7.5M ammonium acetate and 2.5 volumes of absolute alcohol. The samples were shaken and a glass rod was then used to spool out the DNA. Excess alcohol was allowed to evaporate from the DNA before it was transferred to 500µl of TE. DNA samples were stored at 4^0C.

2.3.7 PCR protocol

PCR was carried out on the extracted DNA samples in Omnigene Thermal Cyclers (Hybaid). PCR conditions were 95^0C 5' minutes followed by 30 cycles 95^0C, 57^0C and 68^0C all for a period of 1 minute. The final extension was 68^0C for 10 minutes. The PCR utilised primers within the BLG promoter of the transgene;

BLG FORWARD	5' TCGTGCTTCTGAGCTCTGCAG 3'
BLG REVERSE	5' GCTTCTGGGGTCTACCAGGAAC 3'

The product of these primers is approximately 180bp and can be visualised clearly on a 2.5% TBE agarose gel by electrophoresis.

2.3.8 DNA Digestion and Southern Transfer

DNA from founder mice shown to be transgene positive was analysed by southern blotting to determine transgene copy number. 20μg of DNA was digested with the restriction enzyme Bgl II (New England Biolabs) and the DNA was then transferred onto ZetaProbe™ hybridisation membrane as described elsewhere in this section.

2.3.9 Labelling DNA probes and Hybridisation of Southern Membranes

Transgene copy number was determined using a BamHI/SphI (New England Biolabs) restriction digest product as a probe. DNA probe fragments were gel purified and 20ng labelled with 50μCi, 3000Ci/mol [α-^{32}P]-dCTP (ICN) using High Prime random oligonucleotide premixed solution. (Boehringer Mannheim) This method, first described by Feinberg and Vogelstein (1983) is based on the hybridisation of random short oligonucleotides of 6 bases long to the denatured DNA to be labelled. The complementary DNA strand is synthesised by Klenow polymerase using the 3'OH termini of the random oligonucleotides as primers. Unincorporated nucleotides were removed by passing labelled probes down a Sepahdex G50 column (NICK column – Pharmacia).

Filters were prehybridised in a hybridisation oven (HYBAID) for at least 4 hours at 42°C in UltraHyb (Ambion). Labelled DNA probe was denatured by boiling and then immediately diluted in hybridisation solution overnight at 42°C. After overnight incubation the hybridisation solution was discarded and the filters washed in twice in 2xSSC, 0/1% SDS and twice in 0.1%SSC, 0.1% SDS. The filters were sealed in heat-sealable bags and exposed to BIOMAX MS-1 film in an autoradiography cassette with intensifying screens.

2.4 General Cell Culture Techniques (Teviot Place Protocols)

2.4.1 Maintenance of Embryonic Stem (ES) cells

The ES cell line HM-1 (Magin *et al*, 1992, Selfridge *et al*, 1992) was characterised by shape and size. Cells were small and light refractive when confluent and viewed under light microscopy. ES cells were routinely cultured at 37°C in BHK-21 media supplemented with 1% sodium pyruvate, 1% non-essential amino acids 5% (v/v) foetal calf serum, 5% (v/v) new born calf serum (selected batches), 1.1 ml 50mM β-mercaptoethanol and 1.9 ml LIF (Leukaemia Inhibiting Factor) upon gelatinised tissue culture plates.

2.4.2 Defrosting ES cells

A $25cm^2$ (T25) flask was coated with 3 mls of 0.1% gelatin. The cells were taken from liquid nitrogen and held in a water bath for a few minutes until defrosted. The contents were transferred from the vial into a 15 ml centrifuge tube containing 4 mls of warmed complete + LIF medium. The cells were then spun down at 1000 rpm for 5 minutes at room temperature. The supernatant was removed and the pellet flicked. The cells were resuspended in 10 mls complete + LIF medium and transferred to a T25 flask. They were then placed in a 37°C, 5% CO_2 incubator overnight. The following morning the media was changed on the cells.

2.4.3 Gelatinisation of Tissue Culture Plates

Gelatin (Sigma) was made up to 1% in dH_2O, double autoclaved and stored at 4°C This is equivalent to a 10x stock solution. A 1x (0.1%) Gelatin solution was made and sufficient volume added to cover entire base of tissue culture plate. The gelatin solution was left on the tissue culture plate for a minimum of 30 minutes at room temperature.

2.4.4 Complete Media (CM)

500mls BHK-21 media (Gibco-BRL)

50mls Serum (25mls Foetal Calf Serum: 25mls Newborn Calf Serum)

5mls 100x Non-Essential Amino Acids (Gibco-BRL)

5mls 100x Sodium Pyruvate (Gibco-BRL)

5mls 200mM, 100x L-Glutamine

1.9ml LIF (prepared and titred in lab as per standard protocols – see Smith *et al*, 1988)

1.1ml 50MM β-mercaptoethanol solution (Gibco-BRL)

2.4.5 Trypsinisation of ES Cells

ES cells were routinely passaged 1 in 5 upon reaching approximately 70-80% confluency. Medium was aspirated from cells and the cells washed once in warm PBS. The appropriate volume of (1X) Trypsin-EDTA (Gibco-BRL) was added to cover the entire base of the tissue culture plate. The cells were incubated at 37°C for 2-5 minutes, with occasional agitation of the plate. This provides sufficient time to ensure that ES cells have detached from tissue culture plate and have formed a single cell suspension. ES cell detachment was carefully monitored to avoid over-trypsinisation, as this can result in cell lysis and a reduction in overall ES cell viability. The trypsin was neutralised by addition of a minimum of 2 volumes CM. ES cells were then ready for direct transfer to an area of gelatinised tissue culture plate or for any of the further manipulations detailed in this section.

2.4.6 Freezing ES Cells

Following trypsinisation, ES cells were centrifuged at 1,100 rpm for 3 minutes.
The media was aspirated and the pellet resuspended in an appropriate volume of Freezing Medium (FM).

Plate area Into	Volume of FM	To Fill No. of Cryovials	Area to Re-thaw
24-well	800µl	1	24-well
6-well	800µl	1	6-well
25cm2	1.6mls	2	6-well
75cm2	6.5mls	8	6-well
175cm2	13mls	16	6-well

1ml of FM ES cell resuspension was aliquoted into an appropriate number of cryovials as detailed in the above table and placed in a polystyrene container before being placed in the -80°C as this helps to control the rate of cooling. Cryovials were placed in a -80°C freezer overnight. The following day, the cryovials were transferred to a liquid nitrogen store.

2.4.7 Freeze Media (FM)

80% CM
10% Serum (50% FCS: 50% NCS)
10% DMSO (Di-methyl Sulphoxide)

2.4.8 Electroporation of ES Cells

Following trypsinisation, ES cells were centrifuged in a volume equivalent to 10^8 cells at 1,100 rpm for 3 minutes. The pellet was resuspended in 0.8mls of PBS and transferred to an electroporation cuvette (Bio-Rad) containing 150mg of linearised vector DNA (suspended in 100µl PBS (Gibco-BRL)). The volume of cells was pipetted up and down in the electroporation cuvette once, to ensure good mix of cells with DNA. A cap was placed on the cuvette and pulsed with a charge of 0.8kV, 3µF capacitance, which results in a time constant of 0.1 seconds (Bio-Rad Gene-Pulser). The cells were allowed to recover for 20 minutes before adding 20mls of CM. Cells were plated on 10cm gelatinised plates (Nunc) in non-selective media at an

appropriate density to permit selection the following day ($1x10^6$ per plate for G418 and ganciclovir selection)

2.4.9 G418 Selection and Picking of ES Cell Clones

24 hours after electroporation, the non-selective CM was replaced with selective CM containing 200mg/ml G418 (Gibco-BRL) and 6-thioguanine (6-tG). G418 is a neomycin analogue to which the bacterial *neomycin phosphotransferase* gene confers resistance. Cells which have not incorporated targeting vector DNA (containing a neo cassette), either by random integration or by homologous recombination, will not be resistant to the G418 and will therefore die. Cells which have integrated the targeting vector randomly as opposed to via homologous recombination will express the HPRT selection cassette and as 6-tG is a guanine analogue and substrate for the *Hprt* enzyme it undergoes a phosphoribosylation reaction to generate 6-thioguanine monophosphate (6-thioGMP). High levels of 6-thioGMP within the cells inhibit the biosynthesis of guanine nucleotides and this inhibition eventually results in the death of all cells that express *Hprt* (Miech *et al*, 1967). Selective media on cells was refreshed every 3 or 4 days until colonies of G418 -resistant clones became visible to the eye (usually within 10-12 days post-electroporation). A 96 well plate containing 100ml of trypsin-EDTA in each well was prepared. Selective media was aspirated from the 10cm plate, washed with 1x in PBS and partially aspirated to leave a residual volume of PBS sufficient to cover the base of the tissue culture plate. This measure prevented dehydration of clones. A p200 (yellow) tip was placed over the G418 resistant ES cell clone and gentle suction applied, thus taking the clone up into the pipette. Clones were pipetted into a trypsin-EDTA containing well of a 96-well plate and left for 2-5 minutes at room temperature. The contents of the well were gently pipetted up and down once before transferring to a single gelatinised well of a 24-well plate, containing 1ml of non-selective CM. Non-selective CM media was refreshed the following day.

2.4.10 Thymidine kinase and Selection in Ganciclovir and Picking of ES cell clones

After selection in G418 and 6-tG media to determine clones which have undergone the correct homologous recombination event cells were transiently transfected with Cre protein in order to remove the floxed Neo/TK cassette, hopefully leaving intact a floxed Exon I. This was achieved using transient transfection of a circular plasmid containing the Cre gene (pCre2 plasmid). Transient transfection was achieved using the Tfx-50 lipid preparation (Promega) using the manufacturers protocols. Selection of the correctly targeted clones was carried out using 2μM Ganciclovir to select against the presence of the Thymidine Kinase cassette. CM was prepared which contained Ganciclovir.

2.4.11 Extraction of ES Cell DNA to Enable Subsequent Genetic Analysis

Following trypsinisation, ES cells were centrifuged at 1,100 rpm for 3 minutes. Cells were washed once in PBS and re-centrifuged at 1,100 rpm for a further 3 minutes. PBS was partially aspirated, leaving sufficient residual PBS to cover the pellet. The side of the tube was flicked until the pellet was resuspended in this volume. 600ml of Lysis Buffer was added and incubated at either 55°C for 3 hours, or at 37°C overnight. 500ml of isopropanol was added and shaken on an orbital shaker for 15 minutes, whereby the DNA precipitate became visible. The DNA precipitate was carefully removed using a fine glass hook made from a Pasteur pipette, and washed once in 70% ethanol. DNA was transferred to an Eppendorf microfuge tube containing 50μl TE buffer. Evaporation of any trace residual ethanol and DNA resuspension was achieved by incubation at 60°C for 30 minutes.

2.4.12 Lysis Buffer

100mM Tris-HCl, pH 8.5
5mM EDTA
0.2% SDS
200mM NaCl
100mg Proteinase K per ml

2.5 General Culture of ES cells (NPU Protocols) – Differences in protocols which may affect the outcome of gene targeting

2.5.1 Splitting ES Cells

Cells were fed 2-3 hours prior to splitting. The cells should be around 85-90% confluent at this point. The medium was removed and the cells washed once with PBS. Trypsin-EGTA (TEG) was then added at 1ml (T25), 1.5 ml (T75) or 2 ml (T162). The trypsin was left on the cells for about 1 minute at room temperature. The flask was hit sharply once to dislodge cells and once dislodged medium was added to inactivate the trypsin – 5 ml (T25), 8 ml (T75) or 16 ml (T162). The cell suspension was pipetted quite vigorously to obtain a single cell suspension. The cells were spun down at 1000 rpm for 5 minutes and the supernatant removed. The tube was then flicked to disperse the pellet and the cells resuspended in the required amount of medium. The cells were transferred to a flask and placed in a 37°C, 5% CO_2 incubator.

The flasks or plates should be gelatinised before using. 5 mls 0.1% gelatin per T25 flask, no less than 30 minutes before using. The gelatin is then aspirated before adding medium and cells. Because ES cells divide every 18-24 hours the medium should be changed everyday and the cells passaged every second or third day at a 1:4 split.

Routine change of medium – amounts per flask:

	Weekday	**Weekend**
T25	5-7 mls	10 mls
T75	16-20 mls	25 mls
T162	40 mls	60 mls

2.5.2 NPU Protocol LIF Complete Medium

To a 500 ml bottle of Glasgow MEM (BHK 21) medium (GibcoBRL Life Technologies) the following was added: -

5.7 ml	100 mM sodium pyruvate (GibcoBRL Life Technologies)
5.7 ml	100x non-essential amino acids (GibcoBRL Life Technologies)
28 ml	Foetal calf serum (Globepharm)
28 ml	Newborn calf serum (Globepharm)
1.1 ml	50m M β-mercaptoethanol (GibcoBRL Life Technologies)
4.0 mls	NPU LIF – Batch 2 (7/7/99)

Sterility check all media by placing 10 mls in a sterile container in the tissue culture incubator overnight. Always store at 4°C and use within 1-2 weeks.

2.5.3 NPU Protocol *In Vitro* Cre Targeting

Day 1

Cells: A T75 flask of first round targeted cells was washed and trypsinised and a cell count performed to obtain 10^7 cells. The cells were spun down (1000 rpm, 5 minutes at room temperature) and the supernatant removed. The cells were washed once with 15 mls PBS and spun down again as before. The pellet was resuspended in a final volume of 0.8 mls complete medium (minus serum).

DNA: 25μg of pCre2 plasmid (see Appendix A) was used for each electroporation. Complete medium minus serum was added to make up a total volume of 80μl.

Electroporation: The cells and the DNA were mixed together and transferred to an electroporation cuvette (EquiBio) (0.4 cm electrode, gap 50). The cells were electroporated at 230V (0.23kV) and 500 μFd, pulsed twice (in a BioRad Gene Pulser). This resulted in a time constant of approximately 7.0.

The cells were left to rest at room temperature for 15 minutes before adding complete medium containing LIF. The cells were plated out at 10^4 cells per 10 cm plate. The average number of plates set-up was 40.

Day 6 LIF complete medium containing 2μM ganciclovir was added. The medium was changed as required.

Day 15 Colonies were picked from media-covered plates using a Gilson P200 and a yellow tip and briefly pipetted up and down before being transferred to a 24 well plate containing normal complete media. 50 colonies were selected from the plates and screened for the presence of the targeted allele.

2.5.4 LIF Complete Medium + G418 (250μg/ml)

To a 500 ml bottle of Glasgow MEM (BHK 21) medium (GibcoBRL Life Technologies, Cat # 21710-025) complete media was added

5.7 mls G418 (25 mg/ml) (Promega)

G418 (25 mg/ml) = 0.5g G418 + 20 mls distilled water
This was sterility checked by placing 10 mls in a sterile container in the tissue culture incubator. The G418 media was stored at 4°C and used within one week.

2.5.5 Ganciclovir

The original vial of ganciclovir contained 500 mg (546 mg of Sodium Ganciclovir) of Ganciclovir (MW 255). 5mls of sterile distilled water was added to make a 100 mg/ml solution. A 1:100 dilution in distilled water was made to make a 1 mg/ml working stock solution. A bottle of LIF complete medium contains 569 mls; therefore addition of 290 µl of the 1mg/ml solution of ganciclovir means that the final concentration is 2 µM.

2.5.6 LIF Complete Medium + 2µM Ganciclovir

To a 500 ml bottle of Glasgow MEM (BHK 21) Medium (GibcoBRL Life Technologies) complete media was added:

290µl 1 mg/ml ganciclovir stock solution

This was sterility checked by placing 10 mls in a sterile container in the tissue culture incubator overnight. The media was stored at 4°C and used within one week.

2.5.7 Freezing Cells

Cells were fed one hour prior to freezing down. The cells were washed once with PBS and then 1.5 ml (T75) trypsin-EGTA (TEG) added. This was washed over the cells and left at RT for approximately 2 minutes. The cells were resuspended in 10 mls of complete + LIF medium and then transferred to a 15 ml centrifuge tube and spun down at 1000 rpm for 5 minutes. The supernatant was removed and the pellet of cells flicked gently to dislodge them. The cells were then resuspended in 2.5 mls Freeze Mix A. 2.5 mls Freeze Mix B was then slowly dropped onto the cells and pipetted up and down. The cells were then aliquoted into 1 ml cryotubes (labelled with cell line, passage number and date). They were placed in polystyrene boxes in

the -70° C freezers for at least 24 hours before being transferred to liquid nitrogen for longer-term storage.

FREEZE MIX A	**FREEZE MIX B**
8 mls complete + LIF medium	6 mls complete + LIF medium
2 mls FCS	2 mls FCS
	2 mls DMSO

Freeze mixes should be used cold (from the fridge or on ice). The mixes can be stored at -20°C but must be thawed and mixed well before use.

Flask size	No. vials	Vol Freeze mix A	Vol freeze mix B
T25	1-2	0.5 ml - 1.0 ml	0.5 ml -1.0 ml
T75	4-5	2.0 ml – 2.5 ml	2.0 ml – 2.5 ml
T162	10-12	5.0 ml – 6.0 ml	5.0 ml – 6.0 ml

2.5.8 First Round Targeting of ES cells with a NEOMYCIN cassette

Day 1

Cells: The ES cells were trypsinised and a cell count performed in order to obtain 5 x 10^7 cells. The cells were then spun down at 1000 rpm for 5 minutes at room temperature and the supernatant removed. The cells were washed once with 15 mls of freshly prepared 1x Hepes-phosphate buffered saline (HBS) and spun down again as before. The pellet was resuspended in a final volume of 0.8 mls 1x HBS.

DNA: The targeting construct was linearised and kept in ethanol at -20°C. The DNA was spun down at 13,000 rpm for 5 minutes at room temperature and the supernatant removed. The pellet was washed in 75% ethanol before being spun down again. The pellet was allowed to air dry and resuspended in a final volume of 80 µl 1x HBS.

Electroporation: The cells and the DNA were mixed together and transferred to a cuvette (EquiBio) (0.4 cm electrode, gap 50). The cells were electroporated at 800V (0.8kV) and 3 μFd – the time constant was always approximately 0.1. The cells were then left to rest at room temperature for 15 minutes before adding complete medium containing LIF. They were then plated out at 5 x 10^5 cells per 10 cm plate. 60 plates were set up per experiment.

Day 2

24 hours after electroporation and once the cells have settled down the medium was removed and G418 selection medium added. The medium was then changed every 2-3 days or as required.

Day 15

By day 15 the colonies were now visible to the naked eye. Colonies were viewed under the light microscope to check for morphology and size. Discrete colonies of a medium size and shape were marked for picking. Colonies were picked in 200μl of medium with yellow tips and are transferred to an Eppendorf microfuge tube where they were pipetted gently up and down to disperse the cells. 100μl of the cell suspension was added to a 24 well plate for further analysis and the other 100μl was used for PCR or Southern analysis.

2.5.9 Karyotyping ES cells

Standard methodology was used to karotype ES cell clones. The ES cells to be analysed were grown in a 25cm^2 flask to reach about 50% confluence on the days the cells were to be harvested. 24 hours before the cells were to be harvested, they were fed with fresh medium; a process which is known to cause partial synchronisation. 2 hours before harvesting colchicine was added to a final concentration of 50 μg/ml by diluting into the existing medium (changing the medium completely at this point upsets the partial synchronisation already achieved). Following a 2 hour incubation

at 37°C the medium is removed and kept as mitotic cells tend to be lost into the supernatant. The cells were then trypsinised as usual and the kept medium used to neutralise the reaction. The cells were spun briefly at 1000rpm for 3 minutes and then resuspended in a hypotonic solution (0.075M KCl). The cells were incubated for 4 minutes at room temperature and then spun again. The supernatant was removed leaving only a drop remaining and the cell resuspended in this drop of media. The cells were placed on ice and fixed by drop-wise addition of ice-cold methanol/glacial acetic acid 3:2 v/v. As the fixative was added the tube was flicked sharply. 10mls fixative was added in this way and then the cells spun again and resuspended in 5mls fresh fixative. Again the cells were spun and this time resuspended in 0.5mls of fresh fixative. Drop preparations were made from a height of about 30cm on to a microscope slide washed in 70% ethanol, rinsed in water and removed from the water just before use so that the upper surface remains wet. The slides were held at 45°C to the vertical and three drops of fixed cells dropped onto each slide. The slides were dried thoroughly and the stained with Giemsa stain (Sigma).

2.5.10 10x HBS (Hepes-Phosphate Buffered Saline)

For 200 mls:

192 mM Hepes	10 g
1.37 M NaCl	16 g
50 mM KCl	10 mls of 1M stock
	(=14.91g in 200 mls)
7.1 mM Na_2HPO_4 (dihydrate)	2.8 mls of 0.5M stock
	(=8.9g in 100 mls)
1% Glucose	2 g

All the above reagents were dissolved in a volume of 180 mls of 18.2 mQ water. The pH of this 10x stock solution was then adjusted to pH7.2 using concentrated HCl

and the final volume made up to 200 mls. A 1x HBS solution was made from the 10x stock. The 1x stock becomes pH 7.05 on dilution.

2.6 Animal Breeding

2.6.1 Floxed Stat3 mice

The floxed Stat3 mice have been described before (Chapman *et al*, 1999). Briefly, they were generated by crossing mice with one null *Stat3* allele and one floxed *Stat3* allele (Takeda *et al.* 1998) with mice expressing Cre under the control of the *β-lactoglobulin* milk gene promoter (Selbert *et al.* 1998). Mice were maintained on an outbred background and control BLG-Cre/Stat3$^{flox/-}$ mice were obtained from the same colony segregating for the same combination of genotypes.

2.6.2 BLG-Smad4 Transgenic mice

Founder mice were produced from pronuclear injection (described as in Chapter 3) and were crossed to C57Bl mice to generate stably transmitting F1 progeny. The F1 mice were continuously crossed to C57Bl mice to generate colonies of expressing transgenics.

2.6.3 Harvesting Mammary Glands

Mice were culled by cervical dislocation and an incision made in the midline through the skin but not the body wall. The skin was peeled back on either side and pinned to reveal the mammary glands. The glands were removed and either fixed in formalin for paraffin embedding and histological analysis or snap-frozen in liquid nitrogen in an Eppendorf microfuge tube for protein analysis.

2.7 Mammary Gland analysis

2.7.1 Immunoblotting

Mammary gland was harvested at various timepoints throughout lactation and involution and snap frozen in liquid nitrogen. RIPA lysates were prepared as follows: 50mM Tris HCl, pH 7.5, 150mM NaCl, 1% Nonidet p40, 0.5% sodium deoxycholate, 0.1% SDS. Anti-proteases were added immediately before use; 10μg/ml aprotinin, 1μM pepstatin, 1μM leupeptin, 1mM PMSF and 1mM of the anti-phosphatase sodium vanadate. Tissue was homogenised in RIPA buffer and spun at 13,000rpm. The supernatant was collected and stored at –70°C. Prior to loading, samples were mixed with denaturing sample buffer (NOVEX) and boiled at 95°C for 5 minutes.

Immunoblotting was carried out using NOVEX (Invitrogen) 10% Tris-glycine gels. 20μg of protein was added to an equal volume of 1 x sample buffer containing β-mercaptoethanol. Samples were then boiled for 5 minutes before being spun briefly in a microcentrifuge and loaded on the gel. Gels were run at 125V for two hours or until the markers had separated. Gels were then blotted in half strength Towbins solution (20% methanol, 7.2g glycine, 1.45g Tris in 1L double distilled water). Blots were blocked in milk buffer (TBS, 10% non-fat milk, 0.05% Tris) for 1 hour at room temperature then primary antibodies were added overnight at 4°C. Antibodies used were Smad4, Stat3 and p21 (Santa Cruz), Smad1, phospho-Smad2 and Smad2 (Upstate Biotechnology), Smad7 (gift of Dr C-H Heldin), PAI-1 (Chemicon), Stat1 (NEB). After overnight incubation with primary antibodies blots were washed and secondary anti-rabbit-HRP or anti-goat-HRP (Santa Cruz) added for 1 hour at room temperature. Blots were visualised with ECL (Amersham) and densitometry carried out on three blots from different mice per experiment.

2.7.2 Immunoprecipitation

Immunoprecipitation was carried out as follows; Cell lysates were pre-cleared by adding 100µl of 50% slurry of Sepahrose G beads to 500µl lysate (60µg protein) and incubated for 10 minutes at 4°C. The beads were removed by centrifugation and the supernatant incubated with the primary antibody overnight at 4°C. Antibodies used were p300 (Santa Cruz), Stat3 (Santa Cruz), Smad1 (Upstate Biotechnologies) and Stat1 (NEB). After overnight incubation the immunocomplex was captured by addition of 100µl of 50% slurry of Sepharose G beads by shaking for 1 hour at 4°C. The beads were then collected by centrifugation and washed three times with RIPA buffer before the addition of sample buffer. NOVEX Tris-glycine gels were loaded as described above

2.7.3 Electrophoretic Mobility Shift Assay

Nuclear extracts were prepared from mammary tissue as follows: cells were scraped, washed once with PBS and resuspended in lysis buffer (20 mmol/L HEPES pH 7.9, 1 mmol/L EDTA, 10 mmol/L NaCl, 1 mmol/L dithiothreitol [DTT], 0.4 mmol/L PMSF, 0.1 µg/mL leupeptin, 5 µg/mL antipain and 0.25% (v/v) Nonidet P-40). Cells were incubated in this buffer for 10 minutes before being lysed by 8 to 10 passages through a 26-gauge syringe needle. Nuclei were pelleted by centrifugation at 13,000g for 20 seconds. Proteins were extracted from nuclei by incubation with RIPA buffer at 4°C for 20 minutes with vigorous shaking. Nuclear debris was pelleted by centrifugation at 13,000g for 5 minutes and the supernatant was collected and stored at 80°C.

The nuclear extract was subjected to EMSA analysis as follows:

4µg of nuclear extract was mixed with equal volume binding buffer (50mM NaCl, 40mM HEPES, 3mM $MgCl_2$, 10µg/ml pepstatin, 10µg/ml leupeptin, 15mg/ml

aprotinin, and 1μg salmon sperm DNA) and 0.5ng of ^{32}P-labelled probe. Probes used were SBE and a "mutant" SBE (both Santa Cruz Technologies).

SBE Consensus binding motif	GTCTAGAC
"Mutant" SBE	CATAGCGT

After incubation at room temperature for 30 minutes 0.05% bromophenol blue/xylene was added and loaded onto 5% acrylamide gels and electrophoresed at 150V for 2-3 hours. For antibody supershift analysis 2μl of anti-Smad1, Smad2 or Smad4 antibody (all Santa Cruz Technologies) was added for an additional 20 minutes before electrophoresis.

2.7.4 Histological analysis

Formalin fixed mammary glands were transferred to 70% ethanol before being embedded in paraffin wax. 3μm thick sections were cut and floated onto vectabond-coated slides. Slides were stained with haematoxylin and eosin for histological analysis

2.7.5 Immunohistochemistry

Immunohistochemistry was performed on 3μm thick paraffin embedded mammary gland sections as follows. Smad4 immunohistochemistry was carried out on mammary gland sections from day 10 lactation and day2, 3 and 6 involution samples. Antigen retrieval used citrate buffer (10mM, pH7.6) and the slides were microwaved three times at 700Watts. The slides were then blocked using methanol and the primary antibody was titrated and used at an optimal concentration of 1in 200 overnight at 4°C (goat polyclonal, Santa Cruz). The secondary antibody is a biotin-linked goat anti-rabbit and was used at 1 in 400 (DAKO Corp.). The staining was visualised with an ABC-HRP (horse radish peroxidase) kit (Vectalabs) and then

DAB (3,3 diaminobenzedene) (DAKO), a chromogen which undergoes a chemical reaction leading to brown colouration in positive areas.

2.8 Transfection of ES cells and further analyses

2.8.1 Transfection of ES cells

HM-1 ES cells were grown in conditions as described above. Cells were cultured in 6 well plates until 50-60% confluency. At this point they were either treated by the addition of PBS (1ml 1x, Gibco) or TGF-β (10ng/ml, Sigma) or transfected with one of either the CMV-LacZ reporter construct or 1g of the CMV-Smad1,2,3 and 4 constructs. Transfection was optimised for the Promega TfX-50 transfection system. This is a mixture of a patented synthetic, cationic lipid molecule and L-dioleyol phosphatidylethanolamine (DOPE). The reagent is supplied as a dried lipid film and rehydrated according to the manufacturers instructions. The optimal ratio of Lipid:DNA is 3:1 and once the lipid and expression construct DNA had been vortexed thoroughly they were added to a small volume of serum-free media (0.5mls for a 6-well plate) and left to incubate for 15 minutes before being added to the cells. After 1 hour at 37°C normal media was added to the cells and then they were left until being harvested at the various timepoints. Cells were washed twice and then harvested into PBS by scraping.

Experiments were set up such that transfection efficiency was controlled for by transfection of a CMV-LacZ expression vector (Clontech, US). This vector contains the powerful viral promoter and SV40 polyadenylation site which allows high-level expression of the bacterial *LacZ* gene encoding the β-galactosidase enzyme (Hall *et al*, 1993). β-galactosidase catalyses the hydrolysis of β-galactoside sugars such as lactose. The enzymatic activity in cells can be can be assayed *in situ* with various specialised substrates, which when stained with the histochemical LacZ stain and the compound X-Gal (X-gal is a substrate for the enzyme which gives a blue coloured product) allows quantitation of positively stained cells by eye (Demeneix *et al*,

1994). The procedure consists of washing the LacZ transfected cells once in PBS and then fixing them in 0.05% gluteraldehyde for 5 minutes at 4°C. The cells were then washed twice with PBS, each wash for 5 minutes before the LacZ stain (See **Appendix A**) was added. The plates were then left at 37° to allow the stain to develop before counting positively stained cells.

2.8.2 Annexin V Analysis of transfected ES cells

Harvested cells were washed once in PBS and then resuspended in 100µl of Binding Buffer. To this was added 10µl of fluorescein-conjugated annexin V reagent (R&D Systems, Europe) and 10µl of propidium iodide. The cell mix was gently vortexed and then incubated for 15 minutes in the dark at 20-25°C. Following incubation, and without washing the cells of excess reagents, another 400µl of Binding Buffer (See **Appendix A**) was added to each tube. The cells were analysed on a Coulter Flow cytometer within 1 hour of preparation.

2.8.3 Vindelov Analysis of transfected ES cells

Harvested ES cells were washed once in PBS and spun down in Eppendorf microfuge tubes. The pellet was resuspended in 100µl of citrate buffer and mixed well. To this was added 450µl of solution A and the cells mixed gently by inversion before incubating for 10 minutes at room temperature. After this time 325µl of Solution B was added, mixed and incubated at room temperature and then 250µl of Solution C was added, mixed and incubated on ice for 10 minutes. The samples were then run through a Coulter Flow Cytometer and analysed on EPICS software for DNA content.

Chapter 3 - An *in vitro* analysis of Smad proteins

3.1 Introduction

3.1.1 TGF-β and Apoptosis

Transforming Growth Factor-β (TGF-β) is a multifunctional cytokine which can induce a variety of cellular responses, the most basic being opposing growth effects on cells. This phenomenon depends on what lineage the cells derive from; cells of an epithelial nature are considered to be growth inhibited by TGF-β whereas cells of a mesenchymal nature are generally growth stimulated. The growth inhibition by TGF-β on epithelial cells has been characterised extensively and, depending on the cell system and the experiments themselves, has been shown to be concomitant with apoptosis (Rotello *et al*, 1991 and Bursch *et al*, 1993), induction of CD95/Fas (Ashley *et al*, 1998), down-regulation of c-myc and up-regulation of p27/Kip1 (Kim *et al*, 1998), down-regulation of NF-κB (Azuma *et al*, 1999), down-regulation of bcl-2 (Tsukada *et al*, 1995), down-regulation of p21/Waf1/Cip1 and increased p53 expression (Yan and Sage 1998), induction of Bax (Teramoto *et al*, 1998), induction of a G1 arrest (Motyl *et al*, 1998) and practically every other permutation of responses and genes involved in cell cycle or cell death.

3.1.2 Smads and Apoptosis

Smads are a recently discovered highly conserved group of proteins which have been shown to give specificity to the enormous array of cellular responses to members of the TGF-β superfamily, including epithelial cell apoptosis (Brodin *et al*, 1999, Yanagisawa *et al*, 1998. See section **2.1** of **Chapter 1**). The purpose of this thesis is an examination of the role that members of the Smad family play in an *in vivo* model of epithelial cell apoptosis, namely involution of the mammary gland, a process characterised by proteolytic degradation of extracellular matrix and a loss of secretory epithelial cells by apoptosis. TGF-β has been shown to be important in

many aspects of mammary gland development (Kordon *et al*, 1995, Pierce *et al*, 1993, Jhappan *et al*, 1993) but its role in involution remains to be determined.

Previous work has suggested that the Smad proteins themselves are capable of inducing apoptosis in a variety of cell types including kidney cells (Atfi *et al*, 1997), liver cells (Chen *et al*, 2000) and breast cancer cells (Dai *et al*, 1999) and the first aim of this project was to assess whether a number of different human Smad cDNA expression constructs could induce apoptosis when over-expressed in ES cells.

3.2 Results

3.2.1 Smad Overexpression constructs

The Smad expression constructs were a kind gift of Dr Rick Derynck of the Department of Growth and Development, Programs in Cell Biology and Developmental Biology, University of California at San Francisco, California, USA. The constructs consist of human Smad1, 2, 3 and 4 cDNA sequences cloned into the pRK5 expression plasmid using the CMV promoter as the promoter of choice (Zhang Y *et al*, 1996) (**Figure 3.1**).

E. coli DH5α cells were transformed with 1μg of the plasmid DNA and plated out onto L-amp plates followed by incubation overnight at 37°C. The bacteria were then prepared as described in **Chapter 2** (See **Section 1.1.1**) and the purified plasmid DNA quantified both by Genequant spectrophotometry and on an ethidium bromide plate. All the human Smad expression plasmids were prepared on the same day in the same manner. A control CMV-LacZ plasmid for quantitation of transfection efficiencies was also prepared in the same manner (**Figure 3.2**). The powerful viral CMV promoter and SV40 polyadenylation site allow high-level expression of the bacterial *LacZ* gene which encodes the β-galactosidase enzyme (Hall *et al*, 1993). β-galactosidase catalyses the hydrolysis of β-galactoside sugars such as lactose. The enzymatic activity in cell extracts can then be assayed with various specialised

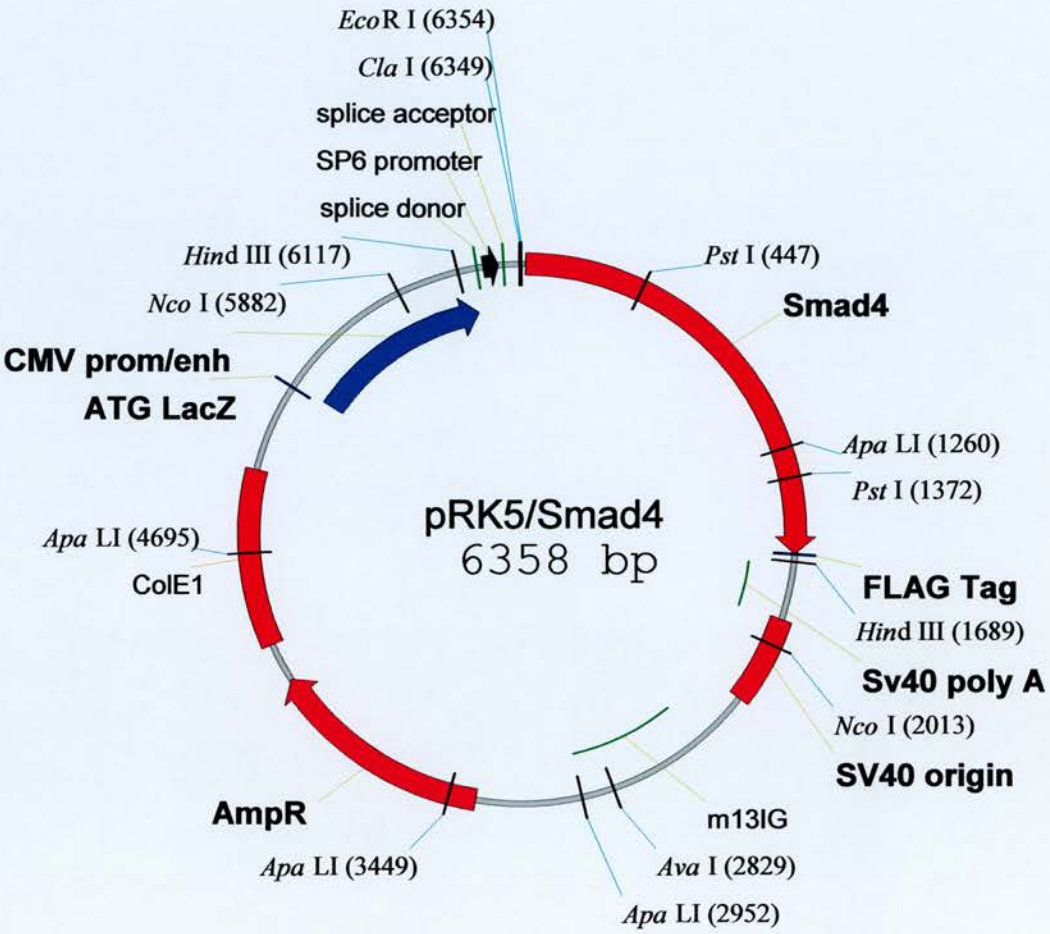


Figure 3.1

CMV-Smad4 plasmid

The CMV-humanSmad4 expression construct consists of the cytomegalovirus (CMV) promoter upstream of a human flag-tagged Smad4 cDNA in the pRK5 plasmid. The Smad4 vector is representative of all the CMV-Smad expression constructs and was a kind gift of Dr Rik Derynck, UCSF, USA (Zhang *et al*, 1996)

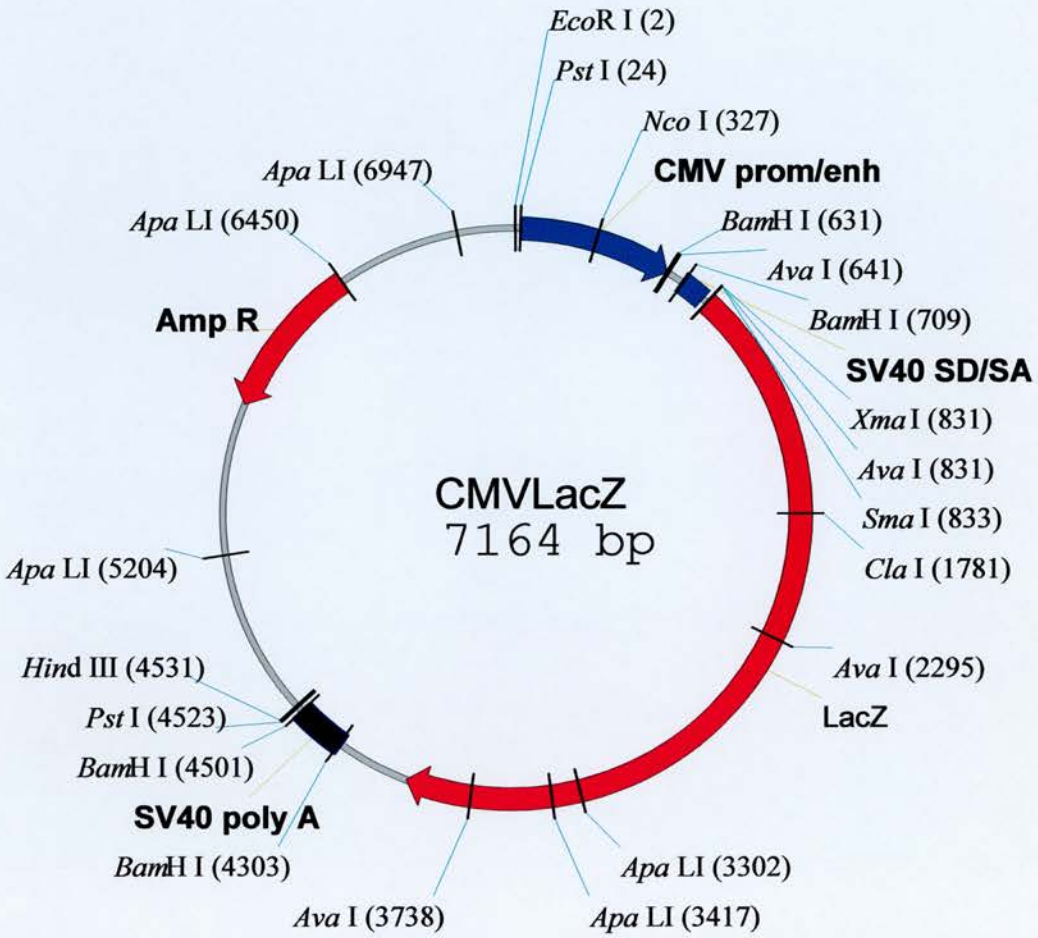


Figure 3.2

CMV-LacZ plasmid

The CMV-LacZ expression plasmid is a commercially available vector (Clontech, US) which allows quantitation of transiently transfected cells via immunohistochemical staining. The vector is an ideal control plasmid for these experiments as it utilises the same powerful viral promoter and SV40 polyadenylation site allowing high-level expression of the bacterial *LacZ* gene which encodes the β-galactosidase enzyme (Hall *et al*, 1993).

substrates, which when stained with LacZ histochemical staining, allows quantitation of positively transfected cells by eye (Demeneix *et al*, 1994). This ability to assay *in situ* expression and therefore transfection efficiency makes the CMV-LacZ plasmid an ideal control.

In order to assess the effect that over-expression of human Smad cDNAs might have on cell viability we chose to over-express the constructs in embryonic stem (ES) cells. There are a number of reasons for this decision; ideally, since the project is concerned with involution and associated epithelial cell apoptosis in the mammary gland, mammary cells would be the model candidate for the *in vitro* work described here. However, mammary gland cell lines often display strange characteristics such as unstable karotype, gene deletions and oestrogen receptor anomalies (Medina *et al*, 1986 and Ormerod and Rudland, 1985). The model solution would be to use primary culture of freshly isolated mammary glands but these are notoriously difficult to transfect and this technique was not available at the time. In a compromise it was decided to use ES cells as they are relatively "normal" (in terms of karyotype and stability) but have one particular phenotype which makes them very interesting for this particular experiment – they do not possess TGF-β type II receptors on their cell surfaces (Goumans *et al*, 1998 and refs therein). This means that the cells cannot respond to external TGF-β via binding of the ligand, although they do have activin and BMP receptors, and thus they are a potentially interesting system in which to examine over-expression of downstream components of the pathway, namely the Smad proteins.

3.2.2 Embryonic Stem Cells

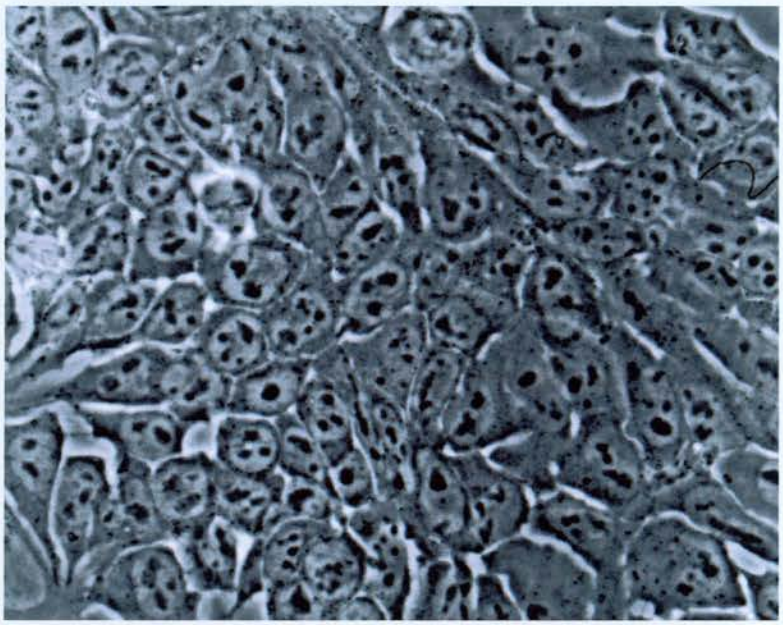
HM-1 ES cells have been described before (Magin *et al*, 1992, Selfridge *et al*, 1992) and are embryonic stem cells derived indirectly from E14 cells (Hooper *et al*, 1987) with a spontaneous deletion of part of the *Hprt* gene (Thompson *et al*, 1989). They display the same morphological features as other embryonic stem cell lines, being cobble shaped cells of approximately 20μm diameter with large nuclei and relatively

little cytoplasm (**Figure 3.3**). Like other ES cell lines they require the addition of Leukaemia Inhibitory Factor or LIF to maintain their pluripotency (Williams *et al*, 1988, Smith *et al*, 1988 and Moreau *et al*, 1988) and this characteristic allows these cells to contribute to the germ line when injected into blastocysts for the purpose of gene targeting experiments.

3.2.3 Annexin V analysis after over-expression of Smads in ES cells

HM-1 cells were grown in the presence of LIF in 6-well plates until they had reached approximately 50-60% confluency. At this time they were then either exposed to 10ng/ml TGF-β or transiently transfected with 1μg of the CMV-LacZ, CMV-Smad3 or CMV-Smad4 expression constructs. Immunohistochemistry with an anti-FLAG antibody was used to show that the transfection of the CMV-Smad expression constructs was possible (**Figure 3.4**). At various time points after this the cells were harvested for Annexin V analysis via flow cytometry. All experiments were carried out in triplicate and repeated on three separate occasions. All experiments were normalised for transfection efficiency against that as determined by LacZ staining in the control wells transfected with the CMV-LacZ plasmid (see **Figure 3.5** for an example of LacZ stained ES cells). Normalisation involved counting the numbers of LacZ stained cells and, by assigning that figure as 100%, percentages of Annexin V positive cells could be made relative to this figure. Annexin V is used as an early indicator of apoptosis since during these early events a loss of membrane asymmetry occurs when phosphatidylserine (PS) is exposed on the outer leaflet of the plasma membrane. Annexin V preferentially binds to PS in the presence of salts and can therefore be used as an accurate indicator of apoptosis (Zhang *et al*, 1997).

At all timepoints tested after treatment with TGF-β the ES cells show very little change in viability and remain intact, the treatment having no discernable effect on the cells (**Figure 3.6**). At 12 hours after transient transfection with the expression constructs CMV-LacZ, CMV-Smad3 and CMV-Smad4 there was a significant increase ($p<0.05$) in the number of Annexin V positive cells to around 10% (resting



(Kindly reproduced with the permission of E. A Aitchison, NPU, Edinburgh)

Figure 3.3

HM-1 ES cells

HM-1 cells are embryonic stem cells derived from E14 cells (Hooper *et al*, 1987) with a spontaneous deletion of part of the *Hprt* gene (Thompson *et al*, 1989). They display the same morphological features as other embryonic stem cell lines, being cobble shaped cells of approximately 20μm diameter with large nuclei and relatively little cytoplasm. Magnification **x 40**.

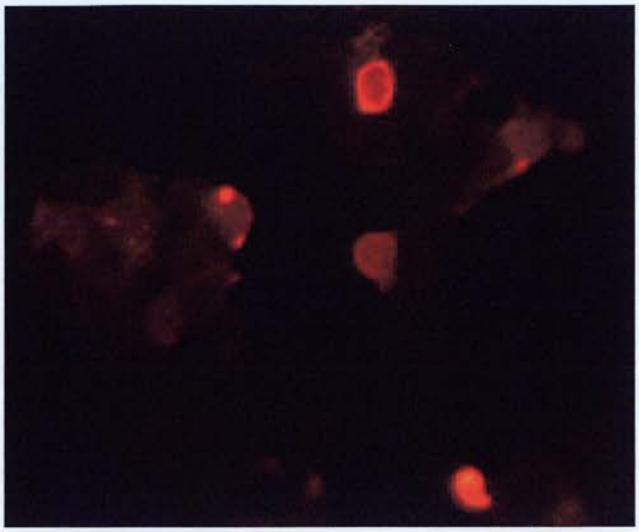


Figure 3.4

TRITC-anti-FLAG immunohistochemistry on CMV-Smad4-flag transfected ES cells

Fixing and staining of Smad4 transfected ES cells shows that most of the staining is localised to the nucleus. Magnification **x 40**.

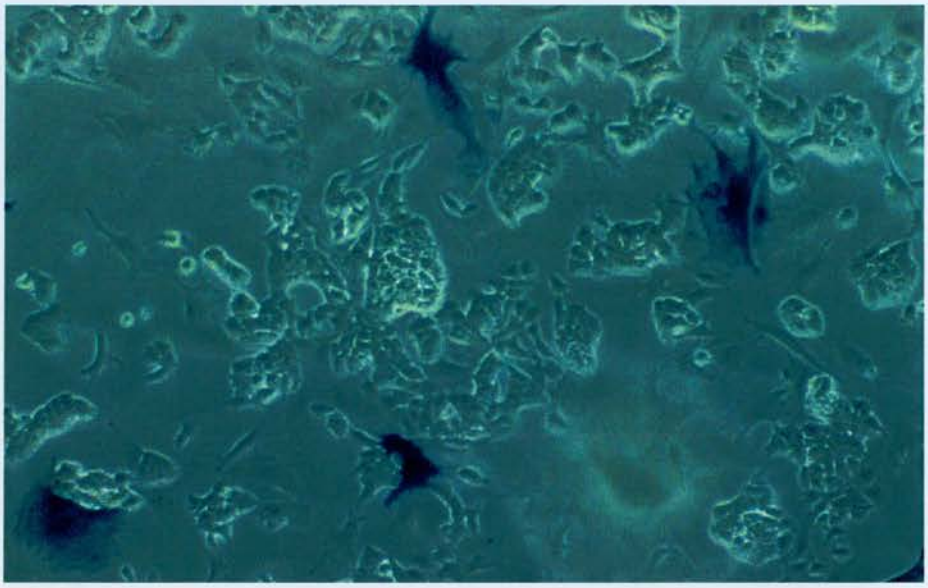


Figure 3.5

LacZ staining of CMV-LacZ transfected ES cells

HM-1 ES cells were transfected with the control CMV-LacZ plasmid and then stained using immunohistochemical techniques to visualise the transfected cells. The overall efficiency of lipid-mediated transient transfection in ES cells is generally quite low – transfection efficiencies of greater than 30% are uncommon. Notably, transfection of differentiated ES cells appeared to be much more efficient. Magnification **x 40.**

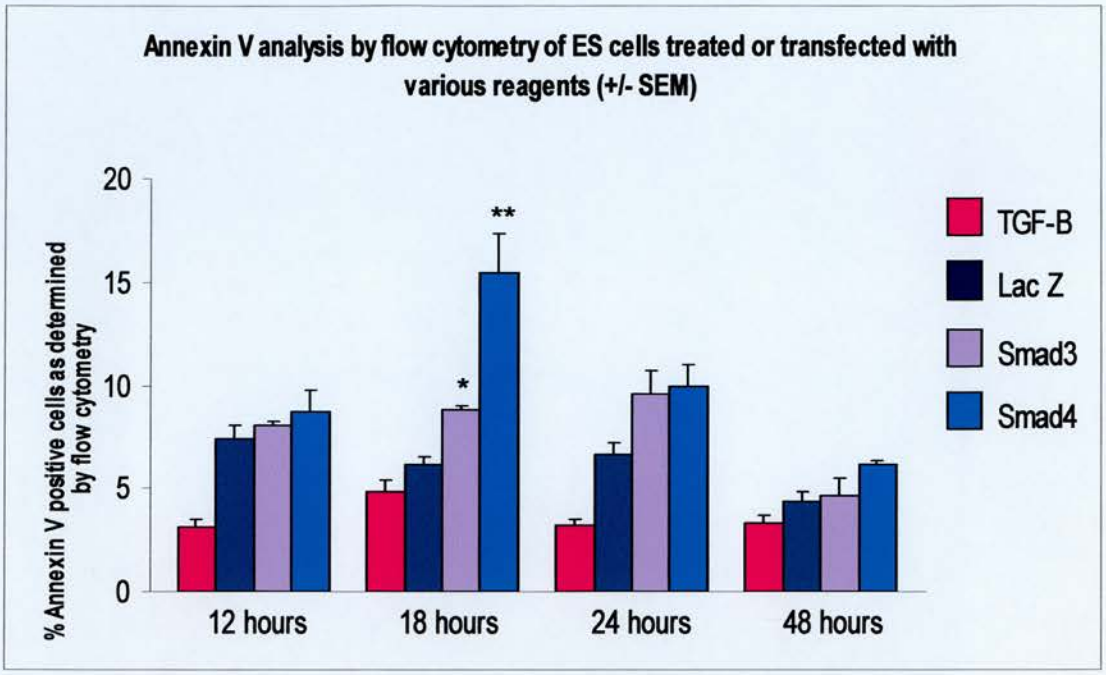


Figure 3.6

Annexin V analysis by flow cytometry of ES cells treated or transfected with various reagents (+/- SEM).

This graph shows the % of Annexin V positive cells after treatment with TGF-β, or after transfection with the control LacZ plasmid or the Smad expression constructs. Annexin V was measured at various timepoints after the treatment/transfection and at 18 hours after transfection with CMV-Smad4 Annexin V levels peaked (~16%) compared to CMV-LacZ (~6%) at the same time point. This increase is significant (P=0.001 using students t-test).

cells display about 4-5% apoptosis) and this is probably indicative of the transfection protocol itself which involves a one hour incubation in serum-free media. At 18 hours after the transfection there was a marked increase in the number of Annexin V positive cells in cultures transfected with the CMV-Smad4 expression construct to a final level of around 16% compared to around 5% after transfection with CMV-LacZ at the same time point ($p=0.001$). Levels of apoptosis induced by Smad3 are also significantly higher than those induced by transfection with the CMV-LacZ control plasmid ($p=0.004$) These levels of apoptosis are not observed when ES cells are transfected with either the control CMV-LacZ plasmid or for the CMV-Smad3 plasmid at this time point.

At 24 hours after transfection with the Smad3 and Smad4 expression constructs, levels of viability had reduced but still remained significantly higher than LacZ induced levels of apoptosis ($p<0.05$) and by 48 hours the transfected expression constructs have no effect on cell viability at all.

3.2.4 Vindelov Analysis after over-expression of Smads in ES cells

In order to determine if this increase in apoptotic cell number was associated with an increase in the number of cells entering G1 (indicative of a G1 arrest) it was decided to analyse cell cycle distribution via a Vindelov analysis (Vindelov *et al*, 1983). This protocol prepares nuclei from cells which can be analysed by DNA content using flow cytometry, thus quantifying the numbers of cells in each of the cell cycle stages. Previous work by Dai *et al* (1999) and Atfi *et al* (1997) has shown that an induction of apoptosis by Smad4 is accompanied by an increase in the number of cells in G1. Having previously shown that treatment of ES cells with TGF-β did not induce apoptosis as determined by Vindelov analysis this study concentrated on the effects of Smad over-expression on cell cycle only.

HM-1 ES cells were grown in the presence of LIF in 6-well plates until they had reached approximately 50-60% confluency. At this point they were then transiently transfected with 1μg of the control CMV-LacZ plasmid or the experimental CMV-

Smad1, CMV-Smad2, CMV-Smad3 or CMV-Smad4 plasmids. Cells were harvested at 18 hours after transfection and nuclei prepared as described in the Materials and Methods section. The results show that 18 hours after transfection with the CMV-Smad4 plasmid approximately 10% more cells have entered G1 phase of the cell cycle (which could be indicative of a partial G1 arrest) than with transfection with either the CMV-LacZ or CMV-Smad3 plasmids ($P>0.05$) (**Figure 3.7**). This result correlates well with the previous published data and supports that concept that Smad4 can induce apoptosis, perhaps through a cell cycle arrest mechanism, as has been shown for TGF-β induced cell death (Bouchard *et al*, 1997 and Mazars *et al*, 1995)

3.3 Discussion

The aim of this work was to test the hypothesis that over-expression of Smads, and especially of Smad4, can cause apoptosis. Although published evidence exists which suggests this is the case these examples have utilised a variety of cell systems and over-expression systems and it seemed prudent to test this hypothesis again using a normal euploid cell, namely ES cells.

This data shows that ES cells do not respond to exogenous TGF-β, presumably because they lack the Type II receptor on the cell surface. This makes an interesting system in which to examine the effects of over-expressing Smads as TGF-β cannot play a role in activating the receptor-regulated Smads2 and 3 via binding to the receptors. However, ES cells do possess activin and BMP receptors and so ligand-binding signals could have conceivably come from these molecules instead.

The results show that Smad4 itself is capable of inducing a statistically significant increase in Annexin V positivity compared with either Smad3 or a CMV-LacZ control plasmid. This shows that over-expression of the Co-Smad, Smad4, can induce apoptosis irrespective of TGF-β bound at the cell surface and suggests that either the receptor activated Smads necessary for Smad4 translocation to the nucleus

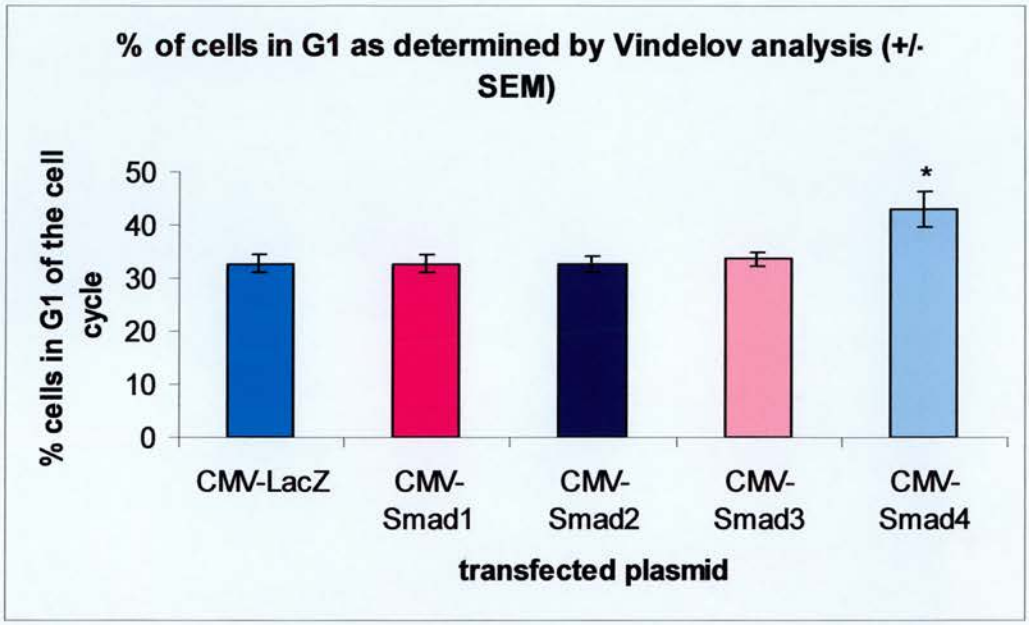


Figure 3.7

% of cells in G1 as determined by Vindelov analysis (+/-SEM)

This graph shows the number of cells in the G1 phase of the cell cycle at 18 hours after transfection with either CMV-LacZ, CMV-Smad1, CMV-Smad2, CMV-Smad3 or CMV-Smad4. The increase in the number of cells in G1 after transfection with CMV-Smad4 (~10%) is significant. $P<0.05$ (students t-test)

have been phosphorylated and therefore activated by another ligand bound at the cell surface (e.g. activin or a BMP) or that they have been activated by some other means currently not known. Either way, over-expression of Smad4 leads to nuclear localisation of the Smad4 protein and subsequently to apoptosis. This data implies that normal levels of Smad4 may be a limiting factor in this progression to apoptosis. Immunohistochemical staining with an anti-FLAG antibody to test that Smad4 does become translocated to the nucleus confirms that nuclear translocation occurs (**Figure 3.4**). This process is only possible if Smad4 was bound to a phosphorylated receptor-activated Smad, such as Smads1, 2, 3 or 5. This scenario seems quite possible

in a cell where TGF-β superfamily signalling is obviously tightly controlled until differentiation when various receptors become highly expressed on the cell surface.

The amount of apoptosis induced by Smad4 over-expression seems relatively small but, although transfection efficiencies between experiments have been normalised, the overall efficiency of lipid-mediated transient transfection in ES cells is generally quite low – transfection efficiencies of greater than 30% are uncommon. Considering then that perhaps only 30% of cells have been transfected, and that these 30% of cells have not been selected for before flow cytometric analysis, the level of apoptosis induced is fairly high and quite specific to Smad4.

The Vindelov analysis is a method of determining what proportion of the cells are in the G1, S and G2 phases of the cell cycle via DNA content on the flow cytometer. This is useful in determining if there is a block in the cycle, which may then lead to apoptotic events, although a direct link of this nature has to date proved difficult to establish. It remains entirely possible that cell cycle arrest and apoptosis could result from parallel effects of the same stimulus. Therefore, it still remains uncertain how cell cycle arrest can lead directly to apoptosis but it is known that TGF-β arrests the cell cycle late in G1 (Laiho *et al*, 1990 and Geng and Weinberg, 1993) and that this is accompanied by maintenance of RB in its underphosphorylated form (Laiho *et al*, 1990) – a phenomenon associated with cellular apoptosis (Dou *et al*, 1995).

The results show that there is a statistically significant increase in the number of cells in G1 of the cell cycle when transfected with the CMV-Smad4 construct. This suggests that Smad transfection leads to a G1 arrest but the true extent of this phenomenon is somewhat debateable. Again the small population of cells transfected was not purified to identify the transfected population prior to flow analysis and consequently the effect seen could be more dramatic if this was the case.

In conclusion I have shown here that overexpression of Smad4 in ES cells results in apoptosis, and that this induction of apoptosis is concomitant with an increase in the number of cells in the G1 compartment of the cell cycle. Future work in this area of the project would definitely include either co-transfection of a GFP expression plasmid with the CMV-Smad plasmids or the generation of a Smad-GFP fusion protein plasmid so that the transfected population of cells can be purified by FACS before Annexin V and Vindelov analysis. Co-transfection in ES cells is notoriously difficult with transfection efficiencies of less then 20% and I would therefore favour generation of the Smad-GFP fusion protein as the approach to follow. Cloning the human Smad cDNAs from the expression vectors into a commercially available GFP fusion protein plasmid should result in an experimental tool which would provide more complete answers to the question of exactly how much apoptosis over-expression of Smad4 can induce in ES cells and clarify whether this is truly associated with a G1 arrest.

Chapter 4 - Generation and Characterisation of a BLG-Smad4 transgenic line

4.1 Introduction

4.1.1 Transgenic Mouse Models

In order to investigate further the role of Smad4 in the mammary gland it seemed appropriate to also attempt a transgenic approach to the problem. Applications involving the phenotypic effects of transgene expression include studies of developmental regulatory genes (Kessel and Gruss, 1990), hormones and receptors (Palmiter *et al,* 1982), oncogenes (Adams and Cory, 1991), genes involved in immune recognition and response (Hanahan, 1989; Goodnow, 1992) and viral genes (Berns, 1991; Skowronski *et al,* 1993). In general, transgenic mice expressing novel genes represent gain-of-function mutations, whereas loss-of-function mutations in many of the same genes can be created only by gene targeting. However, dominant negative mutants have been obtained through expression of transgenes (e.g. collagen; Stacey *et al,* 1988), and in one case, the expression of transgenes encoding antisense RNA has been successful in inhibiting the expression of endogenous genes (Katsuki *et al,* 1988). A detailed description of the history and major advances in mouse transgenic technology can be found in **Chapter 1** of this thesis.

4.1.2 Mouse Models – the Pros and Cons

Mouse mutants provide a powerful but often complex resource with which to understand the functions of genes during development of the mammalian embryo. Although a number of genes which cause developmental disorders in humans have been identified (Muenke, 1995), the underlying molecular and cellular pathways and processes are poorly understood. In the mouse, access to every stage of embryogenesis and beyond is not restricted and, therefore, the mouse has become an important *in vivo* system for studying disease processes and for designing therapeutic

strategies. The third edition of a book which catalogues, amongst many other important aspects of mouse genetics, the naturally occurring mouse mutants, has been published recently (Lyon *et al,* 1996 in *Genetic variants and strains of the laboratory mouse, edn 3.* Oxford: Oxford University Press), indicating the continuing and increasing importance of mouse models. Dense high-resolution genetic maps of the mouse chromosomes are now available as a result of new classes of genetic markers being coupled with the development of new types of mouse crosses for gene and marker assignment (Dietrich *et al,* 1996) and these advances have led to recent successes in using positional cloning strategies to identify some important developmental mutations.

The ability to modify the mouse genome remains unsurpassed when compared with our abilities with other mammalian models. Once candidate disease-related genes are identified, modifying their expression by generating transgenic mice, via either microinjection of mouse eggs or the use of embryonic stem cells, allows their role in disease to be unambiguously deduced. A compendium of artificial gene-targeted mutations has recently been published (Brandon et *al,* 1995, parts I, II and III), reflecting the power and increased use of these techniques. The generation of transgenic mice can allow the identification of new genes not considered previously to play a role in the pathogenesis of a disease. Transgenic mice can also answer more basic questions, such as determining the function of a poorly characterised gene or confirming a give gene function *in vivo.*

The number of transgenic animals now available for study (see http://www.rodentia.com/wmc for a complete listing to date) is an indication that most researchers feel that the benefits of producing such a model greatly outweigh the numerous difficulties associated with this approach. These include problems with germline transmission of the transgenic allele and the effects of uncontrolled copy number and integration site, both of which can lead to unpredicted transgene expression. Because the locus of transgene integration is usually random, the transgene frequently inserts into other functional genetic sequences. Interruption of the normal expression of an endogenous gene may be inconsequential or lethal. The

timescale for generating transgene models is also of consequence, as these experiments typically require one to three years to perform. However, it is generally perceived that transgenic animal systems combine the virtues of cell culture and congenic breeding strategies while avoiding the negative aspects of each system. Using transgenic techniques, a characterised genetic sequence may be evaluated within the specific genomic background of the whole animal. Therefore, transgenic animals may be utilised to study both the function and the regulation of a specific genetic sequence and as such this approach was deemed suitable for an investigation into what role Smad4 plays in the mammary gland.

4.1.3 The Transgenic Approach

I decided to address this question via multiple approaches including both *in vitro* and *in vivo* analyses – the latter consisted of two strategies; a tissue-specific conditional deletion of the Smad4 gene in the mammary gland (see **Chapter 5**) and a tissue specific over-expression of the Smad4 gene via a transgenic model. The aim of this approach was to generate a line of transgenic mice which expressed higher than normal levels of the Smad4 protein during mammary gland lactation and involution. My hypothesis that Smad4 is an essential component of the initiation of the epithelial cell apoptosis associated with involution of the gland suggests that over-expression of Smad4, either before or during this period, may lead to an altered programme of involution.

4.1.4 The Promoter

The first step in the design of a transgenic strain is to select a promoter which will yield strong expression of the gene of interest in the desired target tissue at the desired point in development. This essential component of the transgenic strategy can often be overlooked resulting in an animal with inappropriate expression of the target gene or inefficient expression. One illustration of how promoter differences can influence the outcome of an experiment is given by experiments aimed at generation of mammary gland-specific TGF-β transgenic mice. In 1993 Pierce *et al*

published a TGF-β transgenic mouse where the transgene is driven by the MMTV (mouse mammary tumour virus) promoter (Pierce *et al,* 1993). They showed that expression of the transgene (as confirmed by *in situ* hybridisation, immunohistochemistry and Northern blot analysis) was associated with marked suppression of the normal pattern of mammary ductal tree development in transgenic females. Reduction in total ductal tree volume was observed at 7 weeks, soon after oestrus begins, and was most apparent at 13 weeks, as ductal growth in the normal mammary gland declines. However, during pregnancy, alveolar outgrowths developed from the hypoplastic ductal tree and lactation occurred. Therefore, all the transgenic females could support normal litters. Unlike many other transgenic mouse models in which expression of growth factors or oncogenes under control of the MMTV promoter leads to mammary epithelial hyperplasia and increased tumour formation, the MMTV-TGF-β transgene caused hypoplasia of the mammary ductal tree and no spontaneous tumours were detected in these transgenic animals. In the same year Jhappan *et al* (Jhappan *et al,* 1993) also published studies based on a TGF-β mammary gland transgenic strain which used the porcine TGF-β1 cDNA placed under the control of regulatory elements of the pregnancy-responsive mouse whey-acidic protein (WAP) gene. Females from two of four transgenic lines were unable to lactate due to inhibition of the formation of lobuloalveolar structures and suppression of production of endogenous milk protein. In contrast with the Pierce study, ductal development of the mammary glands was not overtly impaired. There was a complete concordance in transgenic mice between manifestation of the lactation-deficient phenotype and expression of RNA from the WAP/TGF-β1 transgene, which was present at low levels in the virgin gland, but was greatly induced at mid-pregnancy. TGF-β1 was localised to numerous alveoli and to the periductal extracellular matrix in the mammary gland of transgenic females late in pregnancy by immunohistochemical analysis. Glands reconstituted from cultured transgenic mammary epithelial cells duplicated the inhibition of lobuloalveolar development observed *in situ* in the mammary glands of pregnant transgenic mice. These two papers clearly illustrate how promoter differences can definitely influence the outcome of the phenotype.

The expression patterns of MMTV and WAP are quite different, with WAP expression tightly restricted to lactation compared to MMTV expression in the pregnant and lactating gland and in other organs such as the salivary gland (Hennighausen *et al,* 1994 and Muller *et al,* 1988). Although the consequences of using different promoters is clear, it should also be noted that positional effects can greatly influence the outcome of the transgenic phenotype. The classical example of this is when the transgene influences the genome such that different founders from the same experiment can have different phenotypes (Tanaka and Takeuchi, 1991).

4.1.5 Other Considerations

Several other factors are recognised to influence the efficiency of expression of the transgene. This includes the debate about whether introns should or should not be included in the transgene. Experiments have been designed to test the effect of introns on gene expression in transgenic mice (Trudel *et al*, 1987). Four different pairs of gene constructs, which were identical except that one member of each pair lacked all introns, were compared for expression of mRNA after introduction into the murine germ line by microinjection of fertilized eggs. The expression of two chimeric genes, made by fusing either the mouse metallothionein I or the rat elastase 1 promoter/enhancer to the rat growth hormone gene, was assayed in foetal liver or pancreas, respectively, while two natural genes, an oligonucleotide-marked mouse metallothionein I gene and the human beta-globin gene, were assayed in foetal liver. In each case there was, on average, 10- to 100-fold more mRNA produced from the intron-containing construct. Moreover, mRNA levels were proportional to the relative rates of transcription that were measured in isolated nuclei. However, when the expression of the two mouse metallothionein I gene-based constructs was tested after transfection into cultured cells, little difference was observed. These observations suggest that introns play a role in facilitating transcription of microinjected genes although this effect may be manifest only on genes exposed to developmental influences.

The method of introduction of the transgene into the nucleus has also been subject to much testing with microinjection into the male pronucleus now routinely used. In 1985 Brinster *et al* showed that some of the more important parameters that influence successful integration of foreign DNA into mouse chromosomes are the effects of DNA concentration, size, and form (supercoiled vs. linear with a variety of different ends) as well as the site of injection (male pronucleus, female pronucleus, or cytoplasm) and buffer composition (Brinster *et al,* 1985). They showed that the optimal conditions for integration entail injection of a few hundred linear molecules into the male pronucleus of fertilized one-cell eggs and that under these conditions about 25% of the mice that develop inherit one or more copies of the microinjected DNA.

Finally, there has even been debate about the mouse strain used to generate the transgenic animal. Brinster *et al,* (1985) also showed that the overall efficiency depends heavily on the choice of mouse strains; for example, generating transgenic mice that express foreign growth hormone genes is about eight times more efficient with C57/BL6 X SJL hybrid mice than with inbred C57/BL6 mice.

4.1.6 The Beta-Lactoglobulin Promoter (BLG)

In an attempt to circumvent some of the problems described above I decided to use the promoter from the ovine beta-lactoglobulin (BLG) gene as the promoter of choice. This gene has been characterised extensively and its promoter has successfully been used in the past to target gene expression to secretory epithelial cells of the mammary gland (Whitelaw *et al,* 1992, Farini *et al,* 1995, Archibald *et al,* 1990 and Clark *et al,* 1992). The gene encoding the milk protein BLG in sheep is expressed in the mammary gland in a tissue-specific manner during pregnancy and lactation. Expression is induced mid-way through pregnancy, peaks in lactation and begins to decline at the onset of involution (Gaye *et al,* 1986)) The unmodified sheep gene behaves appropriately in transgenic mice, and it has been shown that many of the cis-acting elements that mediate this pattern of expression are located in the proximal 400 bp of the promoter (Clark *et al,* 1998). Expression patterns of the BLG

are also well characterised and it has been demonstrated that BLG-driven transcription is correctly initiated in mice and that BLG-driven synthesis is restricted to the secretory epithelial cells of the mammary gland (Harris S *et al,* 1991).

The BLG promoter has also been illustrated to be less prone to erroneous expression than other mammary gland specific promoters like whey acidic protein (WAP) or MMTV. The major milk whey protein of sheep, beta-lactoglobulin (BLG), is expressed specifically in the mammary gland in a developmentally regulated pattern. Chromatin analysis of isolated sheep nuclei has showed that the promoter resides within a DNAaseI-hypersensitive region in the mammary gland but not in the liver. BLG transgenes display a similar tissue-specific pattern of DNAaseI hypersensitivity in mice and this data has demonstrated an essential role of the proximal DNAaseI-hypersensitive sequences for position-independent expression of the BLG gene (Whitelaw *et al,* 1992).

4.2 Results

4.2.1 Making the Smad4 Transgene

The 4.2kb-BLG/SK+ plasmid (Whitelaw *et al,* 1992) was a kind gift from Dr John Clark at the Roslin Institute in Edinburgh (**Figure 4.1**).

The transgene of choice is the human Smad4 FLAG-tagged cDNA. This kind gift of Dr Rik Derynk (UCSF) seemed an ideal candidate as it was conveniently tagged at the C terminus with the 8 amino acid FLAG peptide sequence, DYKDDDDK, which allows convenient tagging of transgene expression. The FLAG-tagged Smad4 human cDNA in the pRK5 vector is shown in (**Figure 4.2**).

Dr Stefan Selbert devised a cloning strategy and working together we created the BLG-Smad4 transgene. I excised the FLAG-tagged Smad4 cDNA from the pRK5 plasmid with HpaI and ClaI restriction endonucleases. The restriction digest reactions were then run on a 0.7% TBE agarose gel (**Figures 4.3(a) and 4.3(b)**). The

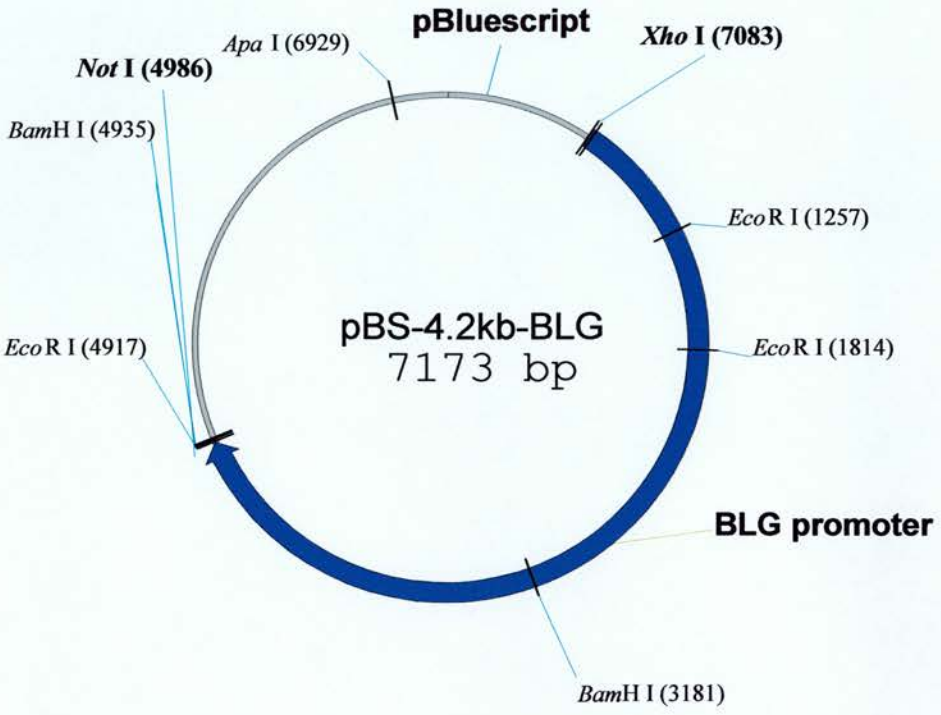


Figure 4.1

4.2kb-BLG/SK+ plasmid

This plasmid contains the 4.2kb ovine beta-lactoglobulin promoter within a pBluescript backbone. This plasmid was a kind gift of Dr John Clark at the Roslin Institute, Roslin, Midlothian

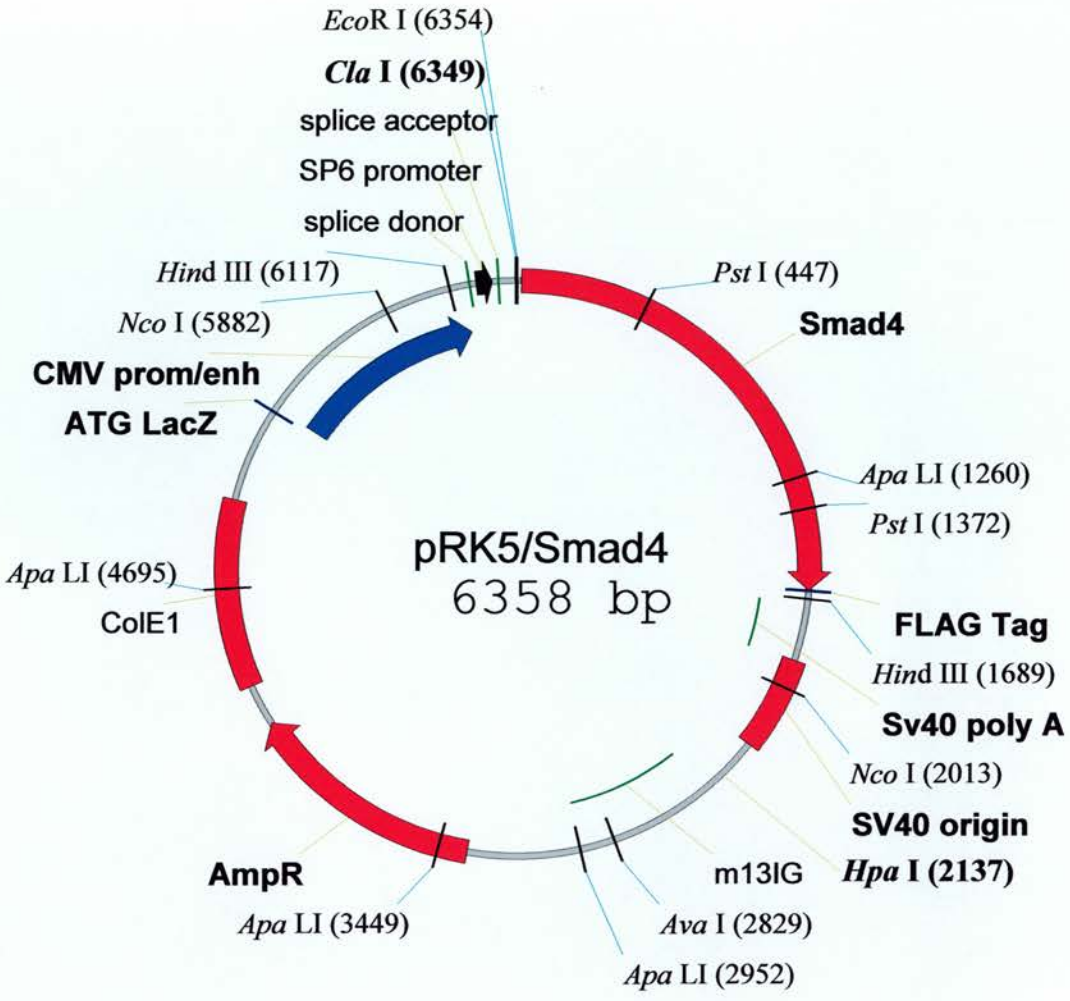


Figure 4.2

pRK5-Smad4-FLAG

This plasmid consists of a FLAG-tagged human Smad4 cDNA positioned downstream of the CMV promoter within the pRK5 backbone. This plasmid was a kind gift of Dr Rik Derynck at UCSF, San Francisco, USA

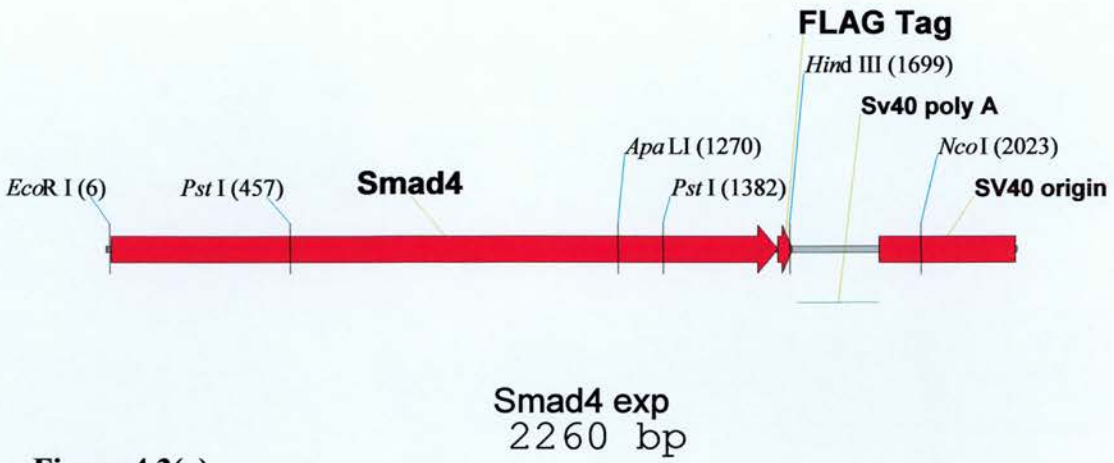


Figure 4.3(a)

FLAG-tagged Smad4 cDNA

This shows the FLAG-tagged Smad4 cDNA schematically once it has been excised from the pRK5 backbone using ClaI and HpaI restriction endonucleases. The fragment is 2260bp in length.

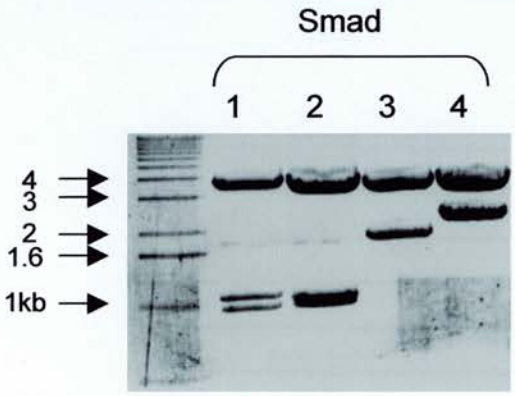


Figure 4.3(b)

Hpa/Cla digest of pRK5 Smads 1-4

This digest shows pRK5-Smad1, Smad2, Smad3 and Smad4 digested with HpaI and Cla I to cut the FLAG-tagged cDNA out of the pRK5 backbone. The digests were then run out on a 0.7% TBE gel.

Smad4 cDNA fragment was then blunted using T4 DNA polymerase and then purified using a QIAEX gel purification kit and 5µl of the purified insert run on a 0.7% TBE agarose gel to check the fragment. (**Figure 4.4(a)**). The 4.2kb BLG/SK+ plasmid has been described elsewhere but consists of the 4.2kb ovine beta-lactoglobulin promoter in the 2.9kb pBluescript SK+ vector (**Figure 4.1** for details). The BLG/SK+ plasmid was cut open with EcoRV restriction endonuclease to leave blunt ends and the purified blunted Smad4 cDNA was ligated using Ready to Go T4 Ligase. An EcoRI restriction digest was performed to determine which clones had the correct orientation of the Smad4 cDNA (**Figure 4.4(b)**). Bacterial clones carrying the correct plasmid (**Figure 4.5**) were grown to saturation and the plasmid harvested (as described in **Chapter 2**) and then prepared for pronuclear injection.

In order to prepare the transgene for pronuclear injection it is necessary to first linearise the transgene. The BLG-Smad4 transgene was removed from the vector backbone by XhoI and NotI restriction digest and then purified by phenol chloroform extraction and ethanol precipitation. A final purification step was carried out using a Qiagen purification kit prior to injection into the pronuclei of fertilised eggs (**Figure 4.6**). The linearised transgene was then injected by Dr Alasdair MacKenzie at the Royal (Dick) Vet School in Edinburgh, UK and by Dr Roberta James at the MRC transgenic unit.

4.2.2 Testing Founder Mice

A number of F1 founder mice were born and these were tail tipped in order to make genomic DNA and the DNA then subjected to PCR to determine the presence of the ovine BLG promoter (**Figure 4.7**). PCR positive mice EK4 and EK6 were mated to C57/Bl mice in order to pass the transgene to the next generation. Founder mice were also tested for the number of integration sites and the transgene copy number by Southern blotting with a BamHI/SphI restriction digest probe against a Bgl II digest. Bgl II cuts once within the Smad4 cDNA (1.6kb downstream from the ATG) and the probe is a restriction digest fragment from the BLG promoter region. The linearised transgene, once cut from the pBS backbone, is

Figure 4.4 (a)

Purified Smad4 cDNA

This figure shows 5µl of purified FLAG-tagged human Smad4 cDNA after digestion from the pRK5 backbone run out on a 0.7% TBE gel. The predicted band size of human Smad4 cDNA is 2.2kb

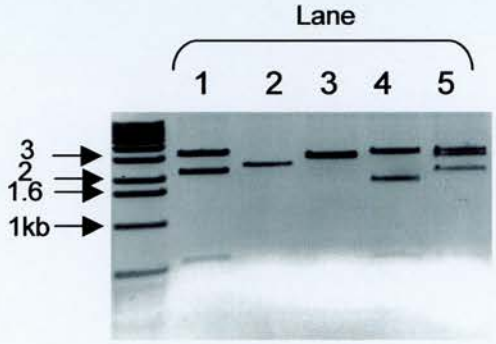


Figure 4.4 (b)

EcoRI digest to confirm Smad4 cDNA in BLG vector

This digest shows a number of EcoRI digests of the cloned Smad4 cDNA in the BLG/SK+ plasmid. The predicted band sizes after EcoRI digest are 3.5kb, 3.1kb, 2.3kb and 600bp. Lane no. 5 has the correct digestion pattern indicative of the correct BLG-Smad4 vector.

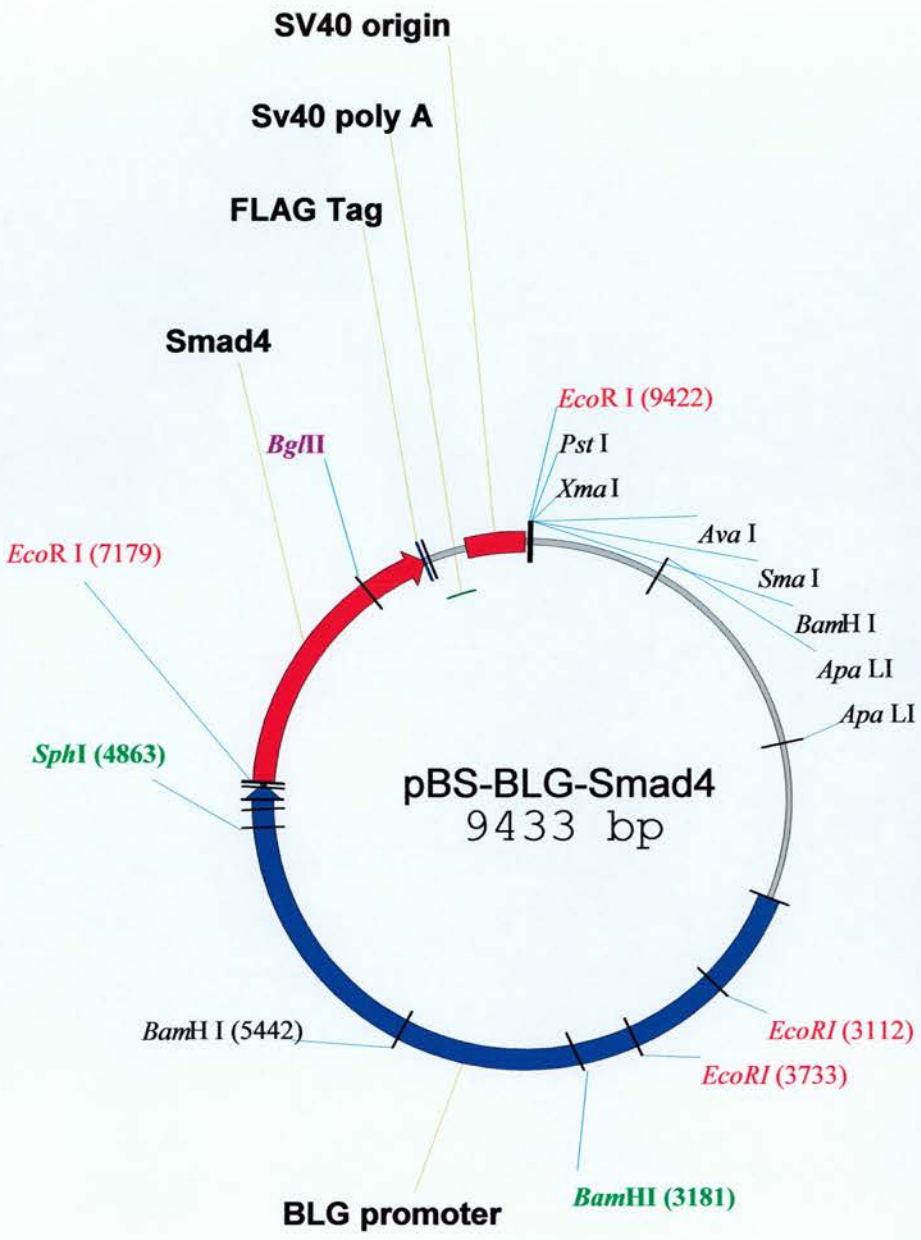


Figure 4.5

The pBS-BLG-Smad4 transgene vector

The complete transgene vector consists of the ovine beta-lactoglobulin promoter upstream of a human FLAG-tagged Smad4 cDNA. Note the BHI/SphI sites which were used to create a probe for testing founder lines by Southern blot. Note also the single Bgl II site within the Smad4 cDNA, used to linearise the transgene for Southern blot analysis.Also outlined are the EcoRI sites (in red) used to detect the correct transgene orientation within the BLG vector backbone.

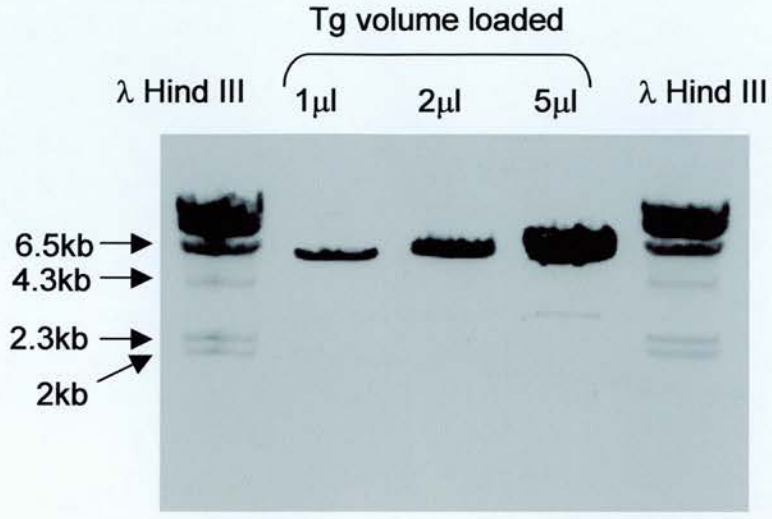


Figure 4.6

Purified Transgene ready for injection

This gel shows the purified transgene after linearisation with digestion with Xho I and Not I. This removes the vector backbone, leaving an intact transgene. The transgene is then purified by phenol chloroform extraction and ethanol precipitation before a final purification with a Qiagen kit just prior to pronuclear injection

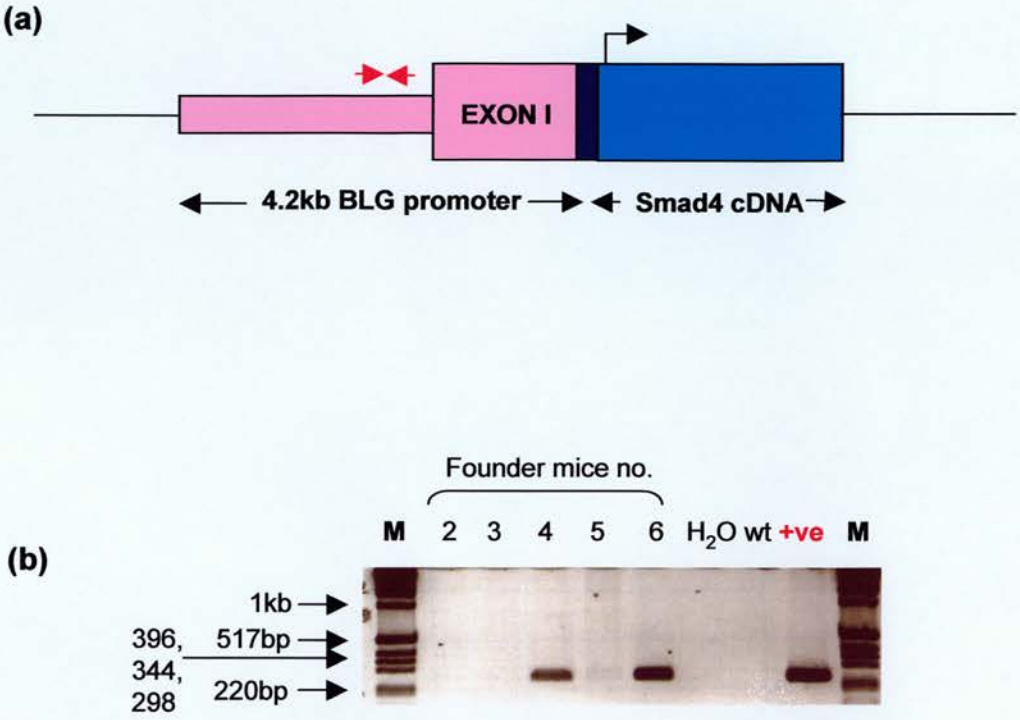


Figure 4.7

Transgene BLG PCR

This gel (**b**) shows the result of the first BLG PCR of the first founder mice obtained by pronuclear injection. The PCR (**a**) utilises primers designed to amplify only a 250bp section of the ovine BLG promoter. Short red arrows represent primers. Founder mice number 4 and number 6 are positive.

approximately 7kb in length. Concatamers of the transgene will produce a 7kb band of varying intensities between founder lines. The number of integration sites can be determined from the other signal present on the autoradiograph as the number of these bands (which may be smaller or larger than the transgene band size depending on each site of integration) is indicative of the number of integration sites. Founder mouse EK6 was tested in this way and was found to have a 7kb band of absolute intensity value of 75.26 and one larger 9kb band with an absolute intensity value of 24.74. Comparison of the intensities between these bands can be used to determine the copy number of the transgene. From this single digest analysis it is likely that EK6 has a copy number of 4 and only one integration site (**Figure 4.8**).

A number of founder mice and wild type mice were also tested for expression of the transgene by the expression of the C-terminal FLAG tag. Western blot analysis showed that none of the PCR positive founder mice expressed the FLAG epitope except for the progeny of EK6: EK62 and EK63 (**Figure 4.9**). The EK6 line was therefore concluded to be the only line which expressed the transgene and the offspring were mated to produce a cohort of both transgene positive and transgene negative animals for further analysis. The timepoints for mammary gland harvesting were day 10 of lactation, day 2 of involution, day 3 of involution and day 6 of involution with the 6 pups per mother being forcibly weaned at day 10 of lactation.

Mammary glands were harvested (two glands each from the upper and lower lateral regions) and immediately frozen on dry ice before being transferred to a -80°C freezer for storage.

4.2.3 Analysis of the Mammary Gland Smad4 Transgenic Mouse Model

A cohort of animals was bred such that a number of animals for each time point and each genotype could be analysed (n=2). Unfortunately time constraints limited the analysis to two mice per genotype per time point. Wild type F1 animals were used as controls for this analysis. The final cohort were genotyped for the presence of the BLG PCR product (**Figure 4.10**) and mammary gland tissue subject to anti-FLAG

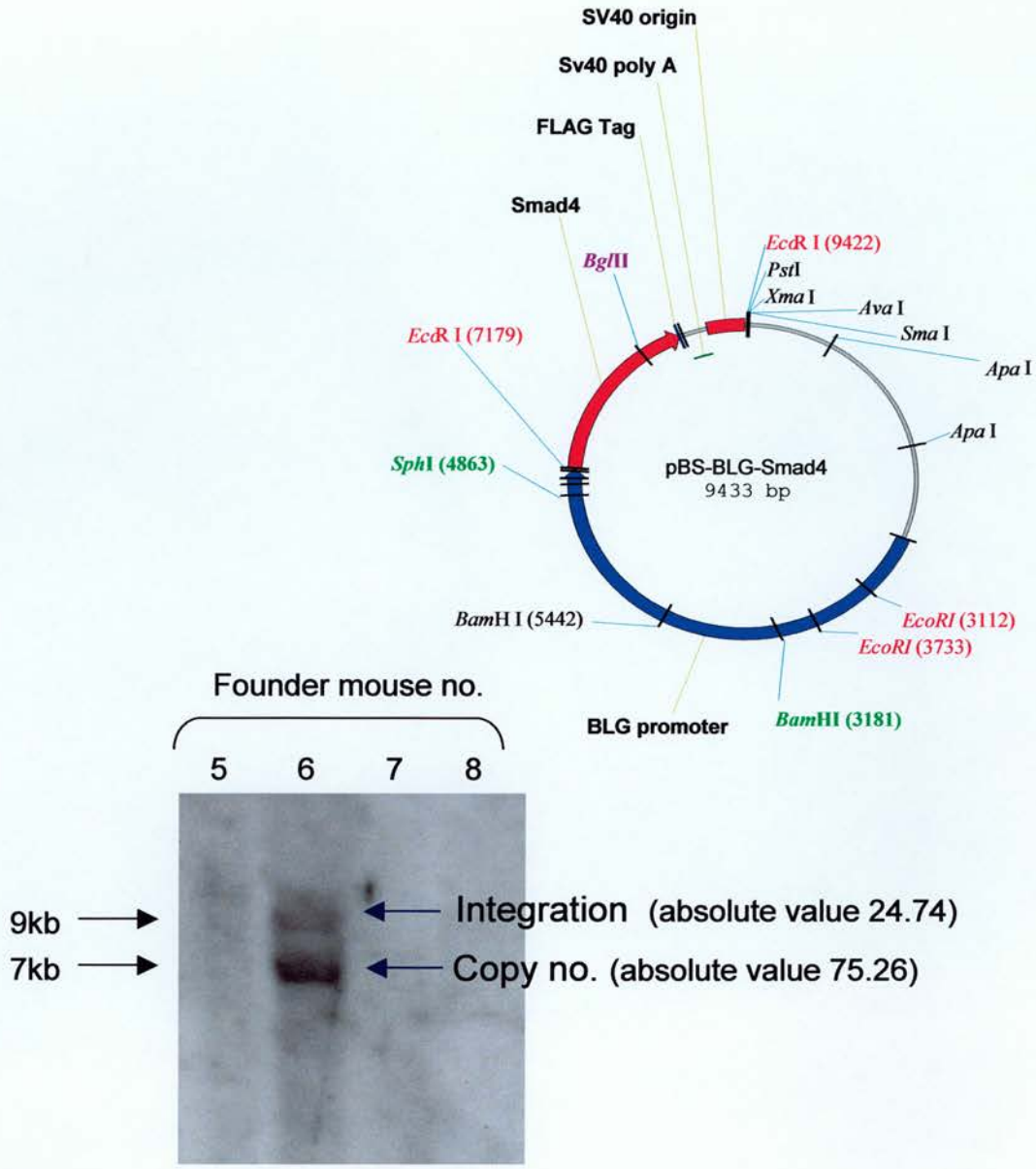


Figure 4.8

Transgene Southern Blot with BamHI/Sph I restriction digest probe on a Bgl II digest

This Southern blot shows the copy number and number of integration sites of the BLG-Smad4 transgene as determined for founder mouse no. 6. The tail tip genomic DNA is cut with Bgl II which cuts once in the Smad4 cDNA. The BamHI/SphI probe is made from a restriction digest within the BLG promoter. The resulting blot shows that founder number 6 probably has a copy number of four and only one integration site (as determined by phosphorimaging – values shown), although another restriction digest would be required to completely confirm this.

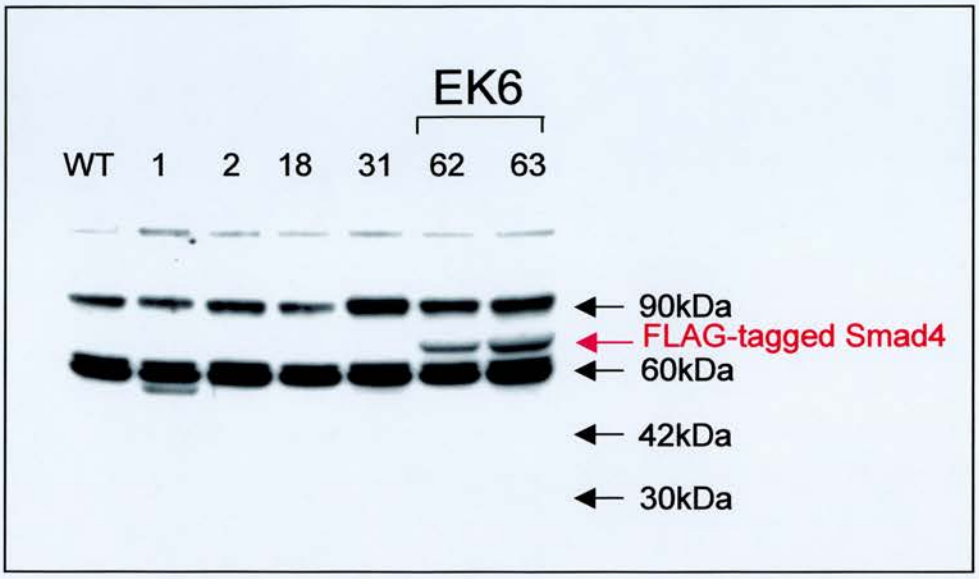


Figure 4.9

Transgene anti-FLAG Western blot

This western blot was performed on whole mammary gland lysates from animals at day 10 of lactation. The blot shows a number of transgene founder mice and F1 progeny mice (EK63 and EK64, both progeny of EK6) after probing with a C-terminal monoclonal anti-FLAG antibody. Only EK63 and EK64 (progeny of EK6) show the FLAG band at just above 60kDa. FLAG-tagged Smad4 should be approximately 60kDa.

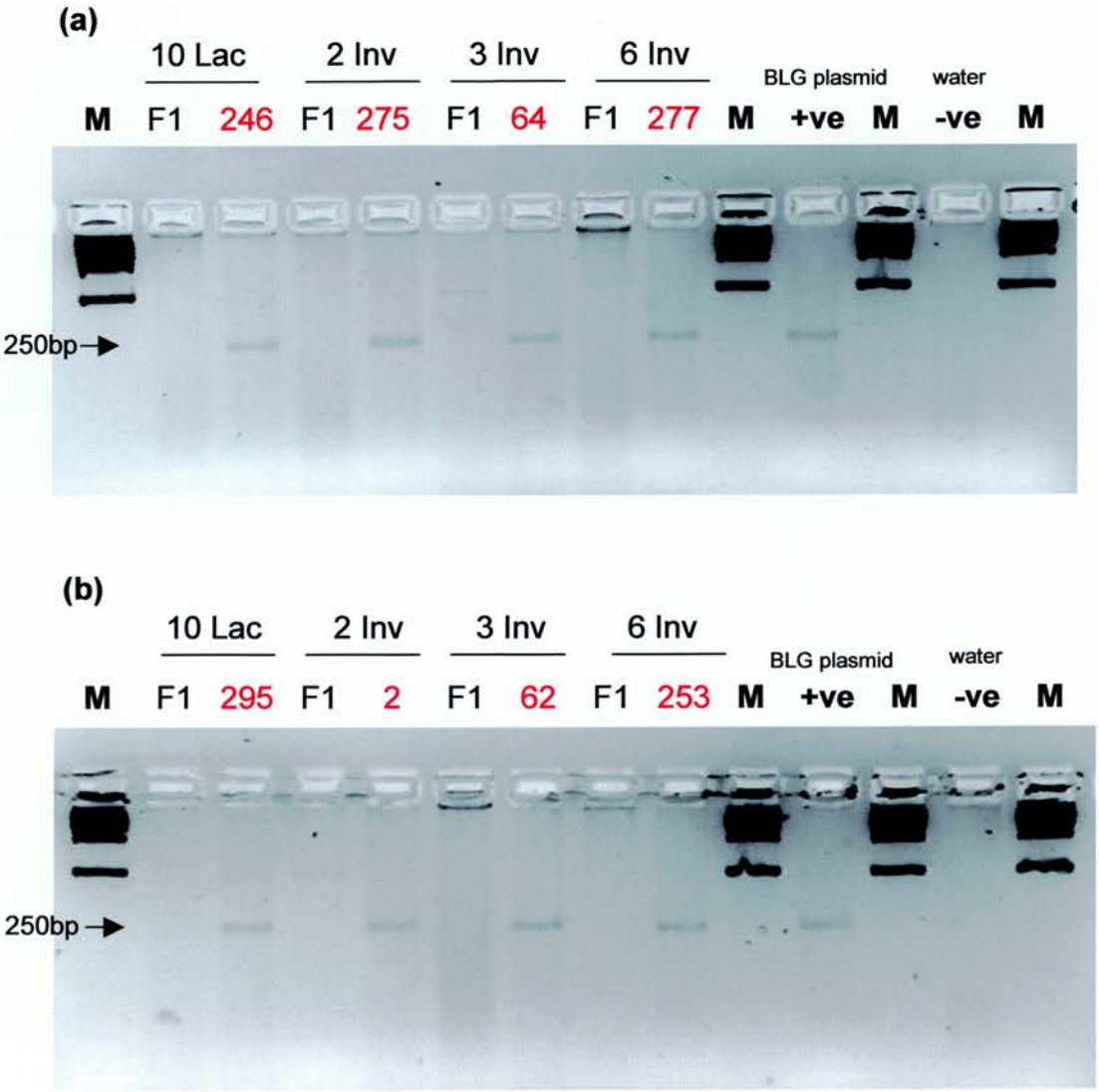


Figure 4.10

BLG PCR on data sets (**a**) and (**b**)

This figure shows the current state of the BLG-Smad4 transgene cohort with a number of transgene positive animals (in red) identified for the various timepoints. At the time of analysis there were limited numbers of transgene negative F2 littermate and therefore F1 mice were used as controls. All the BLG positive transgene mice were tested for FLAG positivity as a marker of the Smad4 transgene expression (See **Figure 4.11**) and preliminary analysis of H&E stained mammary glands has been performed (See **Figure 4.12**).

western blot analysis to determine expression of the transgene (**Figure 4.11**). The FLAG-tagged Smad4 transgene has a predicted size of 60-70kDa.

FLAG expression is seen in all timepoints from day 10 of lactation until day 6 of involution. Some variability was observed between the glands, with EK295 showing slightly less FLAG expression at day 10 of lactation compared to EK246.

H&E stained mammary gland tissues from F1 wild type control animals and BLG-Smad4 transgenic animals were examined for differences in histological phenotype (**Figure 4.12**). It should be stressed at this point that due to time constraints only two mice per genotype, and per timepoint were examined and that this is considered sub-optimal given some expected variance within different mammary glands. Furthermore, time constraints have not permitted a full analysis of the histological data e.g. an assessment of apoptosis rates and an accurate determination of changes in the percentage fat contribution within the gland. However, it is possible to make some general statements about the histology without over interpreting the data. At day 10 of lactation there appears to be little difference between the wild type and transgenic glands although it is possible that the transgenic glands contain slightly more fat than might be expected at this point. By day 2 of involution the wild type gland has begun to show the characteristic degradation of the gland with milk ducts losing their tightly expanded appearance and cell death resulting in epithelial cell loss. In transgenic EK2 this process appears to be slightly accelerated and many of the ductal structures contain sloughed apoptotic fragments associated with this process. In transgenic EK275 the gland seems more intact but there are numerous apoptoses in the lumena of the ducts and in the epithelial layer of the ducts. By day 3 of involution the wild type gland has progressed such that the structure is even less coherent. In transgenic EK64 many of the lumena have collapsed and there is an increase in the amount of fat present in the gland and by day 6 of involution in both the wild type and the transgenics the gland has nearly all collapsed. Both EK253 and EK277 appear to have cleared slightly more of the residual gland than the wild type sample which still retains pockets of parenchymal cells and debris.

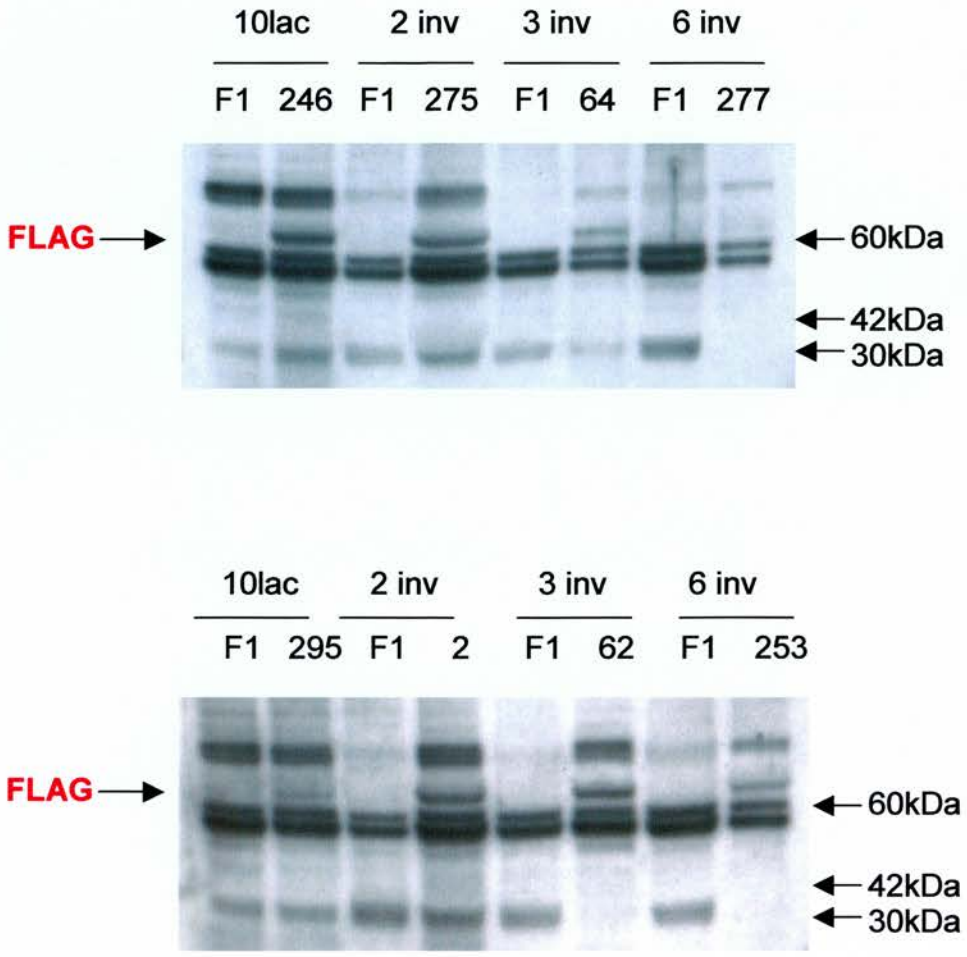


Figure 4.11

anti-FLAG western blots on BLG-Smad4 Transgene positive mice and F1 control mice

The western blots show expression of the FLAG-tagged Smad4 transgene protein only in the EK mice which are positive by PCR for the presence of the BLG promoter. The FLAG tagged Smad4 is about 60-70kDa in size and is present throughout lactation and involution. (30μg protein loaded per well)

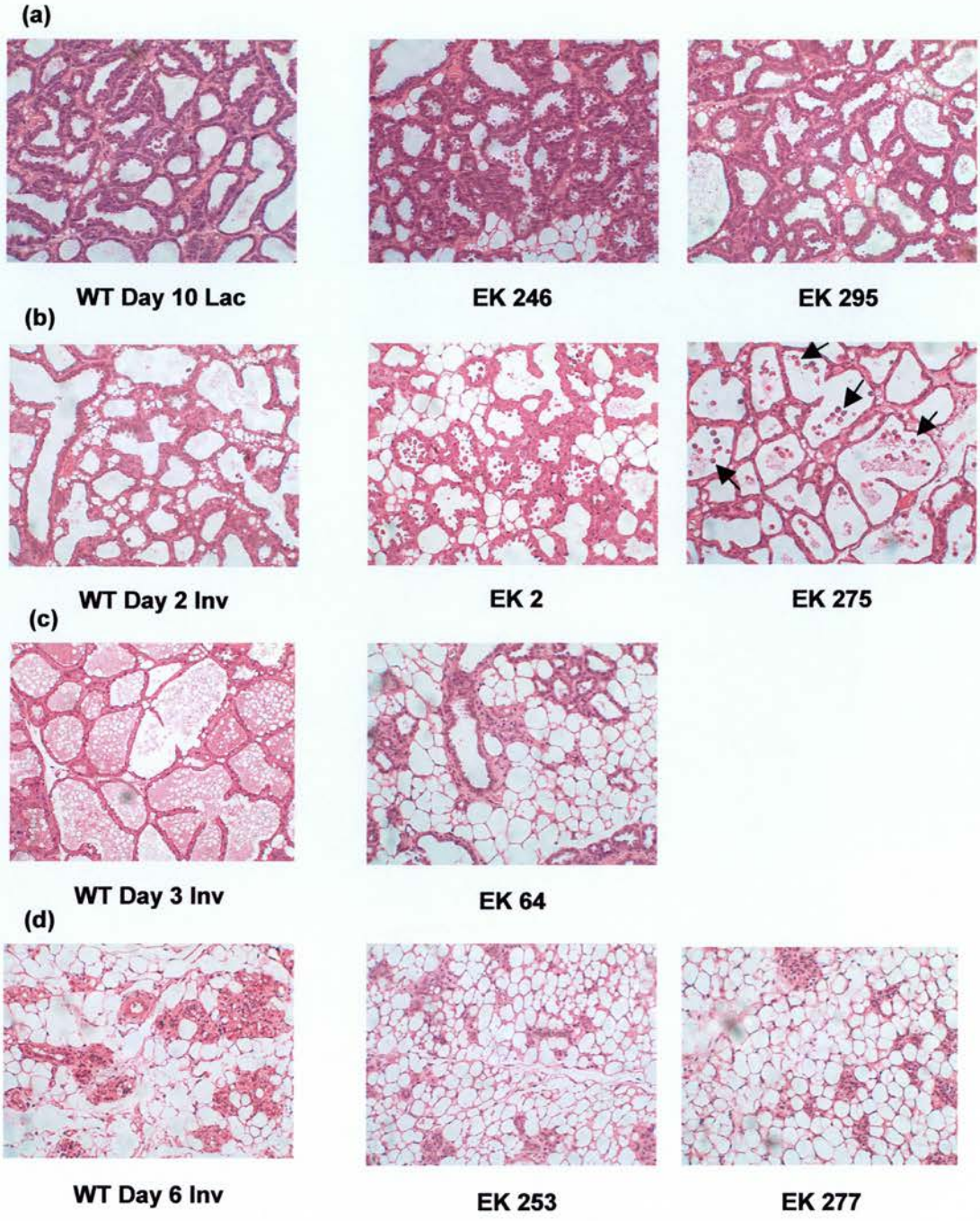


Figure 4.12

H&E stained mammary gland sections from wild type and BLG-Smad4 transgenic mice

These sections show representative areas of mammary gland slides from both wild type and transgene positive mammary glands (**x 40 magnification**). Group (**a**) shows Day 10 lactation, (**b**) Day 2 involution, (**c**) Day 3 involution and (**d**) Day 6 involution. Arrows indicate some of the apoptotic figures.

4.3 Discussion

I have shown here the generation and initial analysis of a BLG-Smad4 transgenic mouse model. The animals expressing the transgene are, at the time of writing, part of a small cohort which cannot be used to produce statistically significant data. However, it is hoped that over the next few months other workers will gather the required littermate control samples and other transgenic samples such that a coherent set of experiments can be performed on a data set with an n = or greater then 3. Once this data has been collected a thorough examination of the phenotype of the BLG-Smad4 transgenic mice can be attempted.

The initial analysis of the H&E stained sections provide preliminary observations which will be confirmed or refuted when the data set is complete. These observations are as follows: throughout involution the transgenic mammary glands differ from the wild type controls in that they show higher levels of apoptosis and generally accelerated involution. Assuming that collection of the final cohort samples is completed there are a number of studies which would be required to be undertaken in order to confirm this phenotype These include a further analysis of the progression of the gland from the H&E slides which would be undertaken using microscopy to count the relative areas of fat per field and number of apoptoses per field. Other markers of gland progression are the matrix-associated proteins such as Matrix Metalloproteases (MMPs) and Tissue inhibitors of Matrix Metalloproteases (TIMPs), and markers of immunological clearance such as the macrophage associated antigen marker CD68. Immunohistochemical stains for these markers could be carried out, as could protein analysis by Western blotting. A biochemical analysis of the tissues will be undertaken to investigate what downstream and related genes are expressed when Smad4 is over expressed in the mammary gland. Western blot analyses will include Smad1, Smad2, phospho-Smad2 and Smad7 levels and an examination of the known downstream markers of Smad activation, p21 and PAI-1. Western blot analysis for markers of apoptosis-associated genes such as the Bcl-2 family will also provide information on what other signal transduction pathways are involved. It will also be interesting to examine protein levels of the STAT family of signal transduction

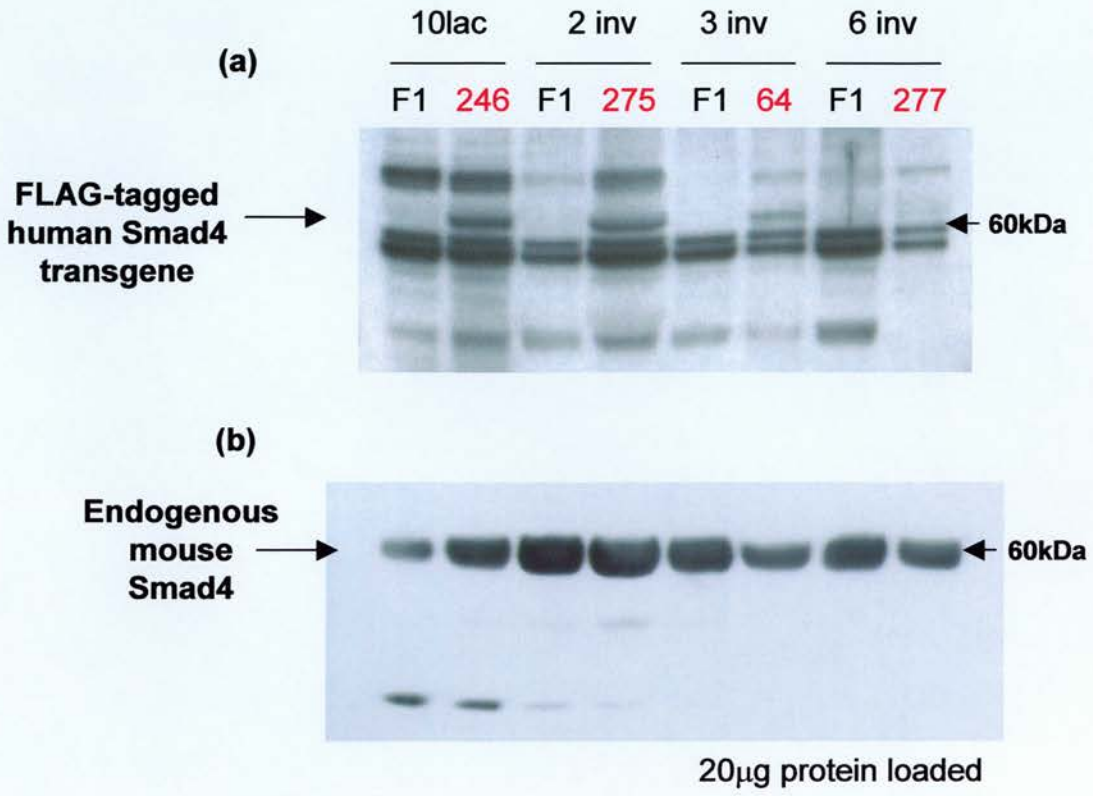


Figure 4.13

Transgene expression levels compared to endogenous expression levels

This figure shows (**a**) human Smad4 cDNA transgene protein levels as determined by Western blotting compared to (**b**) endogenous mouse Smad4 protein levels

molecules as their interaction in Smad signalling in the mammary gland has become apparent during the course of my PhD (see **Chapter 6** for details).

It is useful to note at this point that levels of the transgene expression are perhaps not as high as might be expected and, when compared to endogenous levels of Smad4 in the mammary gland (**Figure 4.13**), it appears that transgene expression levels might be quite low. This is a hard comparison to make for a number of reasons - without the full data set and densitometry analysis of western blot gels it is difficult to assess levels of protein expression, and use of different antibodies to detect FLAG (rabbit polyclonal) and Smad4 (goat polyclonal) makes for very difficult comparisons. However, as only one line of BLG-Smad4 mice have been tested (from EK6) it remains a possibility that this particular line does not express very high levels of the transgene and may therefore not show a particularly strong phenotype. Reasons for poor expression will be discussed in more detail in **Chapter 7** but could include the lack of 3' UTR sequences, position effects and the use of a cDNA transgene as opposed to a genomic construct.

It will be interesting to observe what the final phenotype of these Smad4 mammary gland transgenic animals is. If indeed the preliminary phenotype is confirmed then this will have established an *in vivo* role for Smad4 in initiating apoptosis within the involuting gland, and it is hoped that the information assembled from this particular model will be of help in our understanding of what role Smad4 plays in mammary gland progression.

Chapter 5 - Generation of a Conditional Knockout of Smad4 in the Mammary Gland

5.1 Introduction

5.1.1 Gene Targeting

In order to dissect the functions of individual components of complex biological systems it has become almost routine to engineer some type of transgenic mouse. This approach can yield immense amounts of information regarding the functions of the molecule in question and can also answer questions pertaining to redundancy and compensation. However, there are risks related to the use of this approach and these include not least the time involved to generate the transgene, the requirement for exclusive experimental mouse stocks and then the final analysis. Compared to standard *in vitro* techniques this can therefore be a long term, high-risk approach. However, it is often the only option to fully understand the role that a particular molecule plays in an *in vivo* context and, combined with the other analyses described in this thesis, it was an approach which seemed entirely appropriate for this project. I addressed what role Smad4 plays in the mammary gland in two ways utilising mouse transgenics. First, the generation of a mammary gland transgenic mouse which over-expressed Smad4 (see **Chapter 4** for details) and second, the generation of a "floxed" Smad4 allele. This chapter will deal with the production of the floxed Smad4 mouse.

The ability to disrupt any gene by homologous recombination has led to new approaches in the study of gene function *in vivo*. The derivation from preimplantation embryos of murine embryonic stem cells (ES cells) that retain totipotency even after gene targeting has allowed the generation of intact animals harbouring the desired genetic alteration (Hooper (1992) in "Embryonal Stem Cells"). This technology has been used extensively in recent years to generate "knockout" animals- that is animals in which the desired gene or locus of interest has been ablated from the chromosome, either in a heterozygous or homozygous manner.

These studies have always been of mixed usefulness with some null animals providing great insight into the role that a particular gene plays through exciting or surprising phenotypes for example the TGF-β3 knockout (cleft palate amongst other phenotypes) and the Smad3 knockout (intestinal tumourigenesis) (Kaartinen *et al,* 1995 and Yang *et al,* 1999) and others, simply by a phenotype of embryonic lethality, demonstrating that the gene is of great importance during early embryonic development; like the FGF-4 knockout or certain of the BRCA gene knockouts (Feldman *et al,* 1995 and Ludwig *et al,* 1997). Of course a knockout which results in embryonic lethality is not a failed experiment – it predicts that the gene or protein of interest is extremely important at a certain stage in development and, depending on the lethal phenotype observed, much can be learned from this result - however, it tells little about the gene function in adult tissues.

Often in conventional knockouts other genes will compensate for the null locus resulting in unexpected phenotypes. This phenomenon results from several factors which may complicate the interpretation of a specific phenotype. For example, defects in early development may indirectly also compromise functions in the adult. Conversely, developmental plasticity may compensate for an early loss of gene function. This may occur at the single gene level, the genetic pathway level or involve systemic adaptive mechanisms. Within gene families, related proteins may functionally compensate for deficits associated with the loss of one family member, for example the COX-1 knockout (Reese *et al*, 1999) where COX-2 compensation occurs in the absence of COX-1, and the resulting phenotypes fail to fully address the questions being asked.

5.1.2 Cre Lox Technology

Recently, strategies have evolved which have the potential to overcome this problem by allowing *in vivo* manipulation of DNA in ES cells or animals. This can be achieved through site-specific DNA recombinases; molecules originally described in yeast and bacteria which have the ability to catalyse efficient DNA recombination at sequences of from 25 to 150bp in length, a size sufficiently large that these

recognition sequences are not expected to naturally occur in the mammalian genome. The relative simplicity and efficiency of Cre recombinase from temperate phage P1 has made it particularly useful for this purpose (Sternberg and Hamilton, 1981 and Hoess *et al*, 1982). Placement of recombination sites into the genome and subsequent targeted expression of recombinase have allowed the development of genetic switches that can either ablate or turn on any desired gene in transgenic or gene modified mice.

The Cre-Lox system is described in detail in **Chapter 1** but some of the most important features concerning this application shall be briefly reiterated here.

5.1.2 (a) The Cre-Lox Reaction

Cre is the 38kDa product of the *cre* gene of bacteriophage P1 (Sternberg, 1978 and Sternberg *et al,* 1986) and is a site-specific recombinase of the Int family. Cre recognises a 34-bp site on the P1 genome called *loxP* (locus of crossover of P1) and efficiently catalyses reciprocal conservative DNA recombination between pairs of *loxP* sites (Hoess and Abremski, 1985). The *loxP* site consists of two 13bp inverted repeats flanking an 8bp non-palindromic core region that gives the *loxP* site an overall directionality, that by convention is represented as an arrow (**Figure 5.1 (a i)**). Cre mediated recombination between two directly repeated *loxP* sites results in excision of the DNA between them as a covalently closed circle (**Figure 5.1 (a ii)**). Breaking and joining of the DNA is confined to discrete positions within the core region and proceeds one strand at a time by way of a transient phosphotyrosine DNA-protein linkage with the enzyme. Although the enzymatic details of the early steps involved in these recombination reactions differ, and are in some cases not yet well understood, each case involves a common four-way branched DNA intermediate known as the Holliday junction (Holliday, 1964). A Holliday junction (HJ) is formed when a single strand from each of two duplex DNA segments involved in recombination is exchanged in a region of sequence homology to yield a branched structure that has four duplex arms. The ultimate fate of HJs *in vivo* is cleavage of two of the DNA single strands to yield unbranched products.

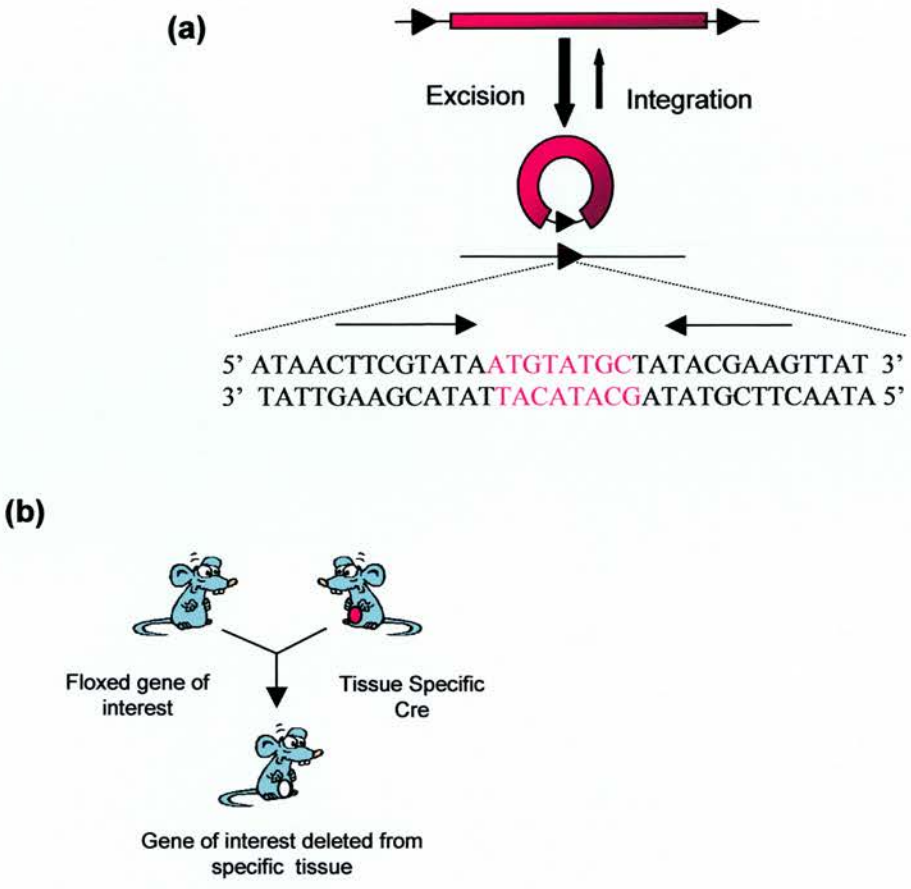


Figure 5.1

Conditional Gene Targeting

Cre recognises a 34 base pair site on the P1 genome called *loxP* (locus of X-over of P1) and efficiently catalyses reciprocal conservative DNA recombination between pairs of *loxP* sites. (**a**) The loxP site consists of two 13 base pair inverted repeats flanking an 8 base pair non palindromic core region that gives the loxP site an overall directionality that, by convention, is depicted as an arrow. Cre mediated recombination between two directly repeated loxP sites results in excision of the DNA between them as a covalently closed circle. (**b**) Correctly targeted ES cells are injected into mice and animals containing the modified gene (which are at this point phenotypically wild type) are then crossed with mice expressing Cre in the desired target tissue and Cre-mediated excision results in a tissue-specific gene ablation

5.1.2 (b) The LoxP sequence

Not all 34bp are essential for efficient recombination; the first 4 bp of either of the 13bp-inverted repeats can tolerate some modification with little if any loss of recombinatorial proficiency or fidelity (Sauer *et al,* 1988 and Sauer 1996). Modified loxP sites might possibly be helpful in the design of genetic switches as the internal 8bp non-palindromic section contains two ATG translational start signals in one orientation (ATGTATGC) but not in the other. It has been hypothesised that it might be important that the loxP sites be oriented such that there remains no false ATG start that could interfere with downstream transgene expression after exclusive recombination. However, there is currently no evidence in support of this, although it has been suggested that loxP sites should be routinely modified such that there is no longer an ATG present. A second potential problem is that because of its palindromic structure, the loxP sequence will form a hairpin structure in transcribed RNA that can, if present in the 5' leader, decrease downstream gene expression (Kozak, 1986 and Sauer and Henderson, 1989). This potential problem can be minimised as already mentioned by introducing a recombinatorially neutral alteration of one or more of the outer 4bp into one of the sites inverted repeats.

5.1.3 Applications of Cre Lox Technology

The exciting potential of the Cre/loxP system became apparent when it was shown that the enzyme would catalyse recombination in the absence of ATP, co-factors or topoisomerase activity (Abremski and Hoess, 1984). Most importantly, the system will also function in eukaryotic cells (Sauer and Henderson, 1988). These observations have subsequently been extended *in vivo* using transgenic mice and it was this work which first demonstrated the potential application of this system to conditional gene targeting (Orban *et al,* 1992 and Lakso *et al,* 1992).

Applications of the Cre-lox system are varied and include cell-restricted, site-specific recombination (as utilised in this project) where regulatory DNA sequences are engineered to be flanked by *loxP* sites which can be introduced into ES cells via

homologous recombination. Once correctly targeted ES cells have been selected (usually via drug selection for an incorporated positive/negative drug selection cassette carrying flanking *loxP* sites) Cre can be transiently expressed in order to remove the selection cassette. These ES cells are then capable of generating heterozygous mouse lines via blastocyst injection which can then be crossed to lines with one null allele to generate mice with either one WT allele and one floxed allele or one null allele and one floxed allele. Crossing these mice to a transgenic mouse in which Cre protein expression is regulated by a tissue or cell-type specific promoter then results in excision of the intervening DNA resulting in a tissue/cell-specific deletion of the DNA of interest (**Figure 5.1 (b)**). This application was first demonstrated by Rajewsky and co-workers (Gu *et al,* 1994) when they deleted the *polβ* exon 1 from thymocytes by using a T-cell specific Cre transgenic via the tyrosine kinase *lck* promoter.

Control of protein expression *in vivo* has also been demonstrated using inducible promoters, although problems such as high basal level of expression, toxicity and poor induction have limited their usefulness to date. Recent modifications of the tetracycline-responsive regulatory system now provide an effective inducible promoter *in vivo* (Hennighausen *et al,* 1995 and St-Onge *et al,* 1996). Combining site-specific Cre-lox recombination with inducible promoters offers perhaps the ultimate flexibility in designing experiments to generate conditional alterations, not only in a particular tissue/cell type but also at varying times, overcoming the increasingly problematic limited availability of tissue and stage-specific promoters.

The elegant potential of the Cre-lox system was chosen as an attractive approach to examining the role of Smad4 in the mammary gland for two main reasons; conventional knockout technology at the Smad4 locus had resulted in embryonic lethality (Sirard *et al,* 1998) and mammary gland β-lactoglobulin-Cre transgenic mice were readily available which have been well characterised (Selbert *et al,* 1998) and would allow a thorough investigation into the role that Smad4 might play in the epithelial cells of the mammary gland.

5.2 Results

5.2.1 Screening the λ library

In order to generate a targeting vector for use in homologous recombination in embryonic stem cells (ES), it was necessary to first isolate a genomic clone that contains part if not the entire murine Smad4 locus. The initial part of this work (the cloning and majority of the characterisation of the genomic clone) was carried out together with Dr Stefan Selbert at the Department of Pathology, Teviot Place, Edinburgh University. It was decided to isolate a murine Smad4 genomic clone from the mouse genomic λPS library (MoBiTec), a eukaryotic genomic library created in the vector system λPS. This system was chosen as molecular cloning of a given locus not only requires the screening of the library, but also the subsequent characterisation of the phage recombinants, and as this latter stage can be quite complex (purification, restriction mapping and subcloning) this method was selected as it circumvents many of these problems. It features an automatic plasmid subcloning facility of insert fragments and therefore combines the advantages of large insert genomic libraries with the convenience of working with high copy plasmids (**Figure 5.2**). λPS is a replacement vector of lambda origin which can accommodate DNA inserts of up to 20kb. This allows the complete representation of complex genomes in a reasonable number of recombinants. The linear vector contains two loxP sites in direct orientation flanking high copy plasmid backbone and the insert. Recombination between these two sites is mediated by Cre recombinase, leading to excision of the multi-copy plasmid from the phage genome. Hence, this is termed automatic subcloning. It can be achieved by transferring the phage into the Cre-expressing *E. coli* strain BNN132.

The library was plated at $2x10^6$ pfus total and filter hybridisation was carried out using a previously generated murine Smad4 cDNA probe.

I created a cDNA probe using the following primers:

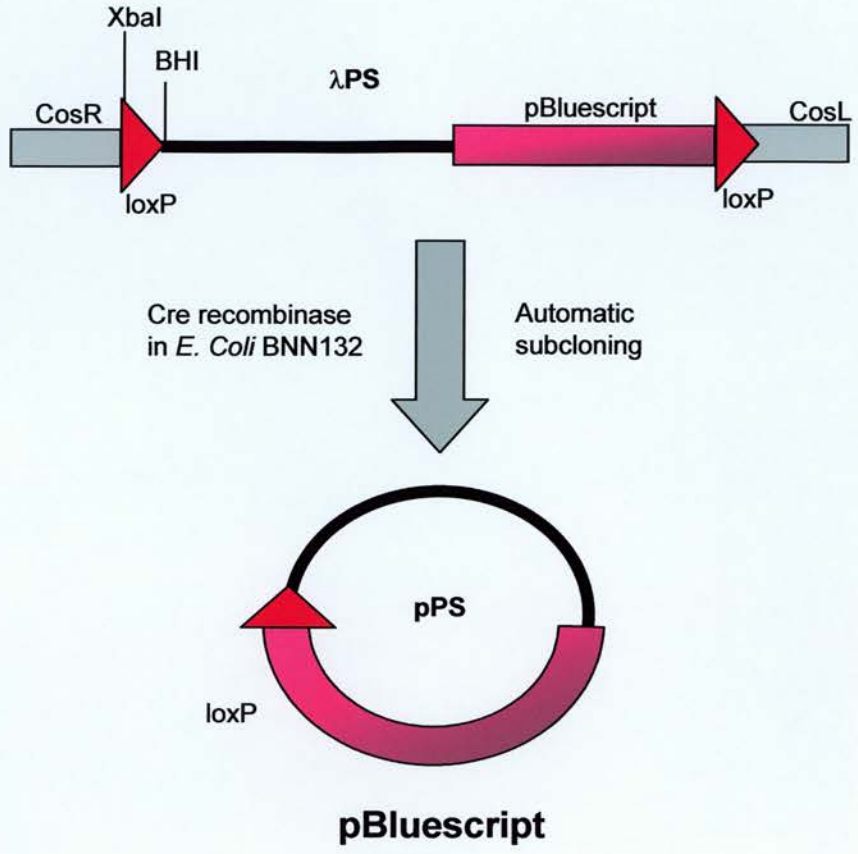


Figure 5.2

Structure of the λPS vector

λPS is a replacement vector which can accommodate DNA inserts of up to 20kb. The linear vector contains two loxP sites in direct orientation flanking a high copy plasmid backbone and the insert. Recombination between these two sites is mediated by Cre recombinase, leading to excision of the multi-copy plasmid from the phage genome.

DPC4 (1)	5'-GATGTGTCATAGACAAGGTGG-3'
DPC4 (2)	5'-CAATTCCAGGTGAGACAACCC-3'

These primers were designed to amplify the murine MH1 region of the Smad4 gene and gave a product of approximately 2kb (**Figure 5.3**).

Positive plaques were purified by additional rounds (usually 3) of plating and hybridisation (**Figure 5.4**) and purified phage added to the Cre-containing BNN132 cells, resulting in automatic subcloning of the positive genomic clone.

The result of screening the library was a 15kb genomic clone which was then mapped extensively to identify suitable sites for the insertion of the single loxP site and the floxed Neo/TK cassette. The mapping of the 15kb genomic clone consisted of a number of approaches including single and double restriction digests, Southern blot hybridisation to the digested DNA using PCR probes and dideoxynucleotide chain termination sequencing into the DNA .

I designed primers based on the known human cDNA sequence of DPC4 (Genbank Accesssion No. U44378) to the MH1 region (5'), the proline-rich linker region and the MH2 region (3') as follows:

DPC4 (1)	5'-GATGTGTCATAGACAAGGTGG-3'	**MH1**
DPC4 (2)	5'-CAATTCCAGGTGAGACAACCC-3'	
DPC4 (3)	5'-GTTCACGACTTTGAAGGACAGC-3'	**Linker**
DPC4 (4)	5'-CTGAGCAGTAAATCCATTCTGC-3'	
DPC4 (5)	5'-GCTCCATTGCTTACTTTGAAATG-3'	**MH2**
DPC4 (6)	5'-TTGGGTAGATCTTATGAAACTGC-3'	

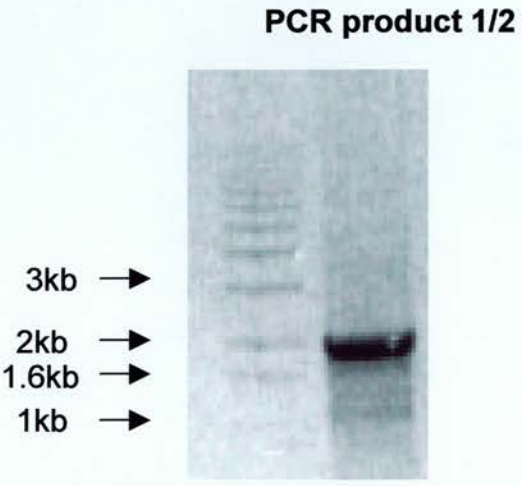


Figure 5.3

DPC4 1/2 PCR product

Primers 1 and 2 amplify the 5' MH1 domain of the Smad4 gene. This product is approximately 2 kb in size and was used as a probe for screening the λ library.

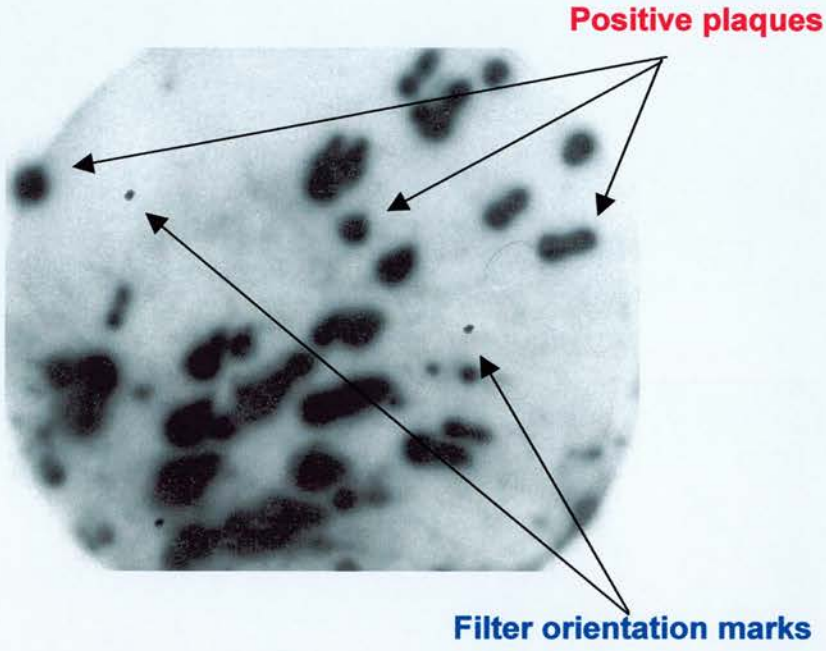


Figure 5.4

Typical filter from screening the λ library

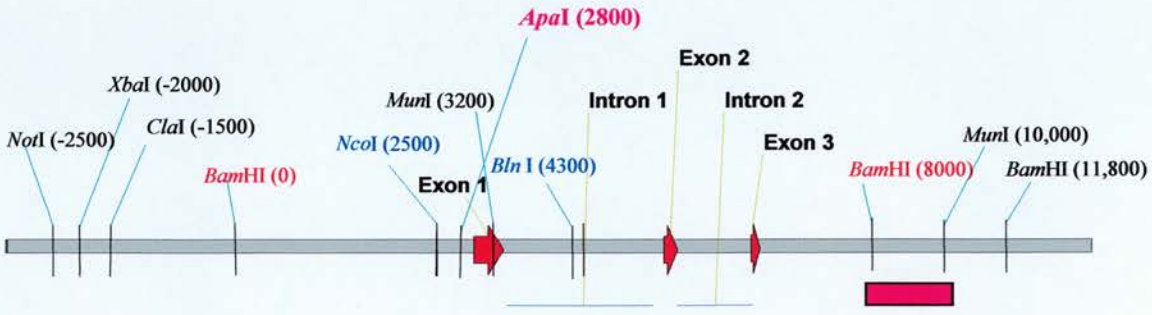
This photograph shows positive plaques after hybridisation of the filter with the Smad4 PCR product **1/2**. A number of positive plaques can be seen. The positive plaques were picked and screened in subsequent rounds.

5.2.2 Creating the floxed Smad4-HPRT targeting vector

Using these techniques the 15kb genomic λ clone was mapped (**Figure 5.5**) such that it was possible to isolate an 8kb BamHI fragment and this was then subcloned into the MCS of pBluescript II (KS+) (Genbank Accession No. X52327, Stratagene) for further analysis (**Figure 5.6 and Figure 5.7**). This 8kb fragment was also subjected to restriction digest and then probed with the Smad4 cDNA probes described above. Southern blot analysis confirmed that the 8kb genomic clone contained only the 5' MHI region of the Smad4 gene and sequencing from the pBluescript vector into the 8kb clone confirmed the presence of exon III at the 3' of the DNA and exon I at the 5' end. It was therefore used as the basis of the floxed targeting construct as it was found to contain exons 1, 2 and 3 of the murine Smad4 gene (**Figure 5.7**).

The floxed Neomycin/TK cassette was a kind gift of Professor Reinhard Fässler and has been described before (Potocnik *et al*, 2000). Briefly this cassette consists of a Neomycin antibiotic gene and a Thymidine Kinase gene with single loxP sites on either side of the cassette in the same orientation. The floxed Neo/TK cassette can be excised from its pBluescript II (KS+) backbone by restriction digest with BamHI and XhoI (**Figure 5.8**).

The floxed Neo/TK cassette was cloned into the single Bln site of the 8kb fragment which had already been determined by sequencing as being approximately 1kb downstream of exon I and about 1kb upstream from exon II of the murine Smad4 gene (**Figure 5.9**). A single loxP fragment (**Figure 5.10**) was then cloned into the Nco I site of the 8kb genomic clone, a site determined to be 700bp upstream of exon I. LoxP fragments were engineered to have EcoRV and BamHI restriction sites at either end of the 34bp palindromic loxP sequence in order to aid in cloning and to help determine orientation of the loxPs by restriction digest. The final step in generating the targeting construct was the introduction of an HPRT cassette into the NotI site of pBluescript i.e. outside the 3' region of homology. The rationale behind this is that if the targeting construct were to integrate randomly (that is not through homologous



15kbλ clone restriction digest map

Figure 5.5

15kbλ clone restriction digest map

This figure shows a map of the 15kbλ clone as determined by restriction digest, southern blot and sequencing. The BamHI sites which were used as the basis of the 8kb subcloned fragment are outlined in red text. The 2.2kb BamHI/Mun I fragment used as a probe for 3' southern blots is outlined as a pink bar. The map also shows the sites intended for insertion of the single loxP sites (NcoI site outlined in blue) and the floxed Neo/TK cassette (Bln I site outlined in blue). The Apa I site used in the 3' Southern blot strategy to detect correctly targeted ES cell clones is shown outlined in pink.

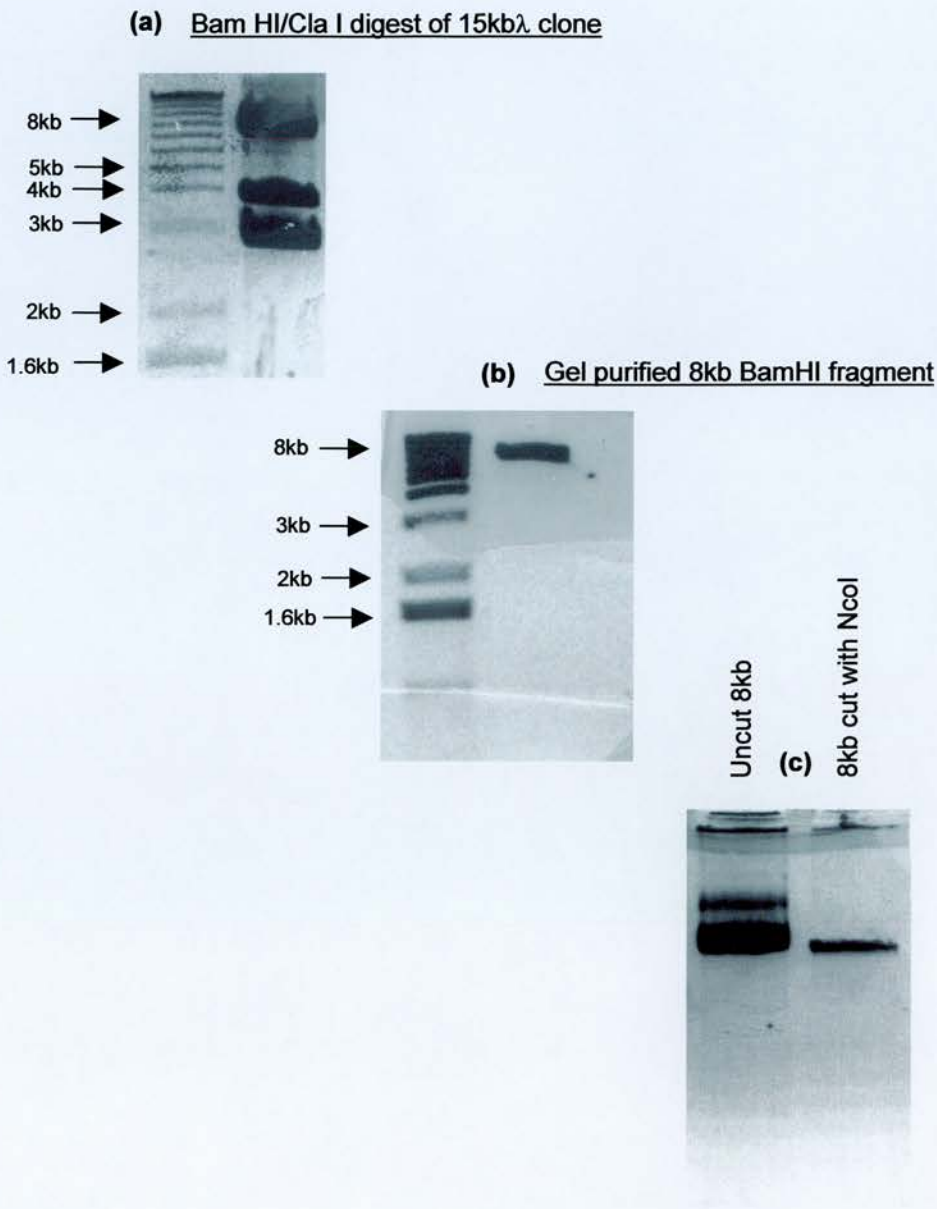


Figure 5.6

Subcloning the 8kb BamHI fragment from the 15kbλ clone

These figures show (**a**) BHI/ClaI digest of the 15kb λ clone showing that there is an 8kb fragment. Figure (**b**) shows the purified 8kb band before it was subcloned into pBluescript. Figure (**c**) shows uncut 8kb plasmid and plasmid cut with NcoI I. This single site within the 8kb BHI fragment was used to insert a single loxP.

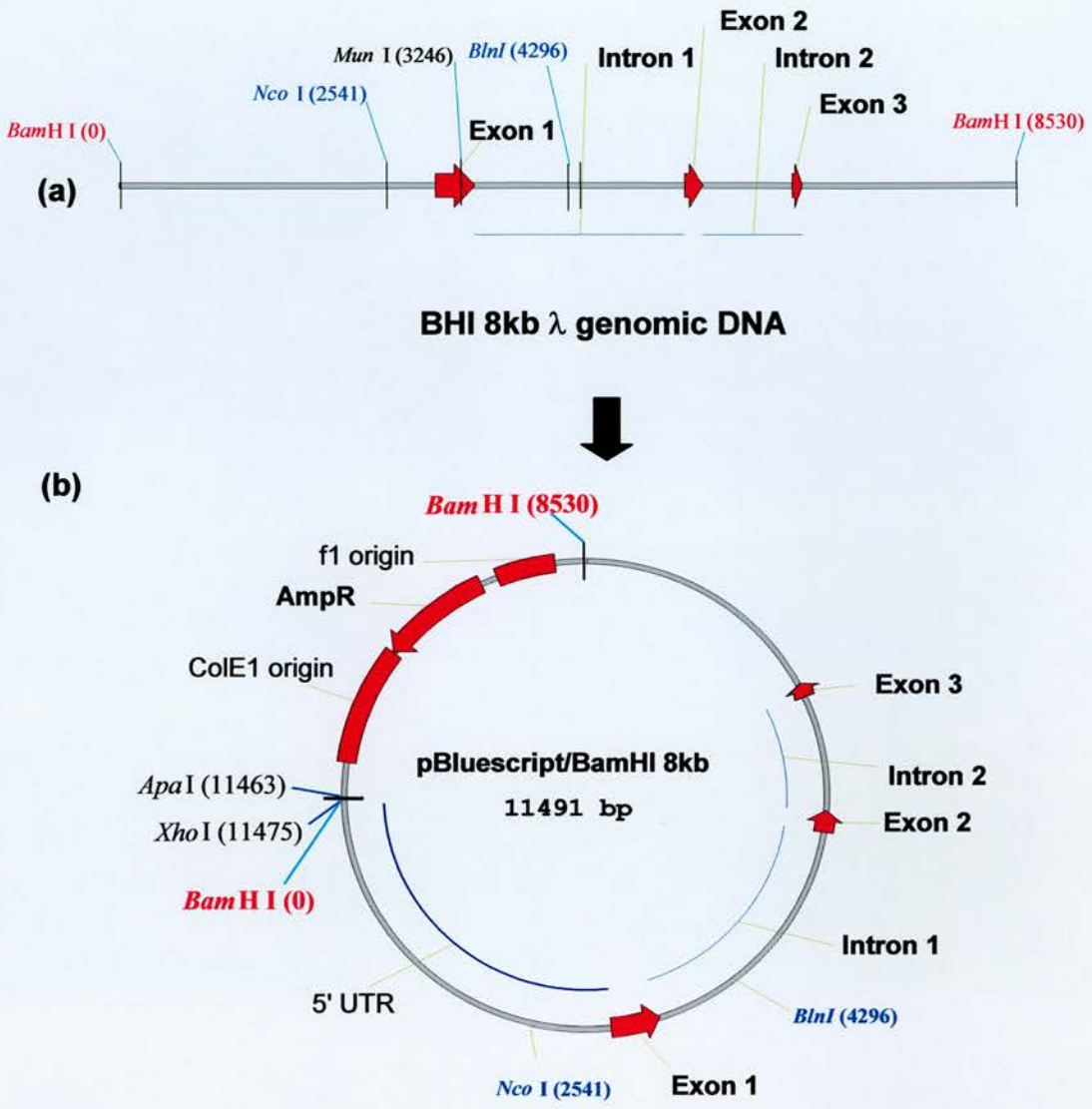


Figure 5.7

The BHI 8kb λ clone

(**a**) An 8kb BHI fragment was sub cloned into pBluescript from the original 15kb λ clone. The 8kb clone was mapped and found to contain Exons I, II and III of the Smad4 gene. This 8kb fragment then became the basis of the targeting construct. The cloning strategy was devised by Dr Stefan Selbert and involved insertion of the single loxP site into the NcoI site upstream of Exon I and insertion of the floxed Neo/TK cassette into the BlnI site between Exons I and II, thus placing loxP sites on either side of Exon I. (**b**) This figure shows the sub cloned 8kb λ fragment containing Exons 1, 2 and 3 of the murine Smad4 gene in pBluescript.

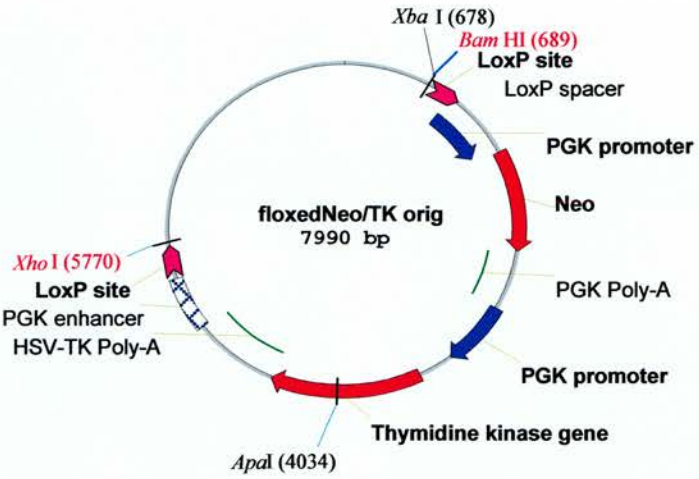


Figure 5.8

Floxed Neo/TK cassette vector

This figure shows the floxed Neo/TK cassette (a kind gift of Professor Reinhard Fässler) before it was cut out of the vector backbone and ligated into the Bln I site between Exon I and Exon II of the Smad4 gene.

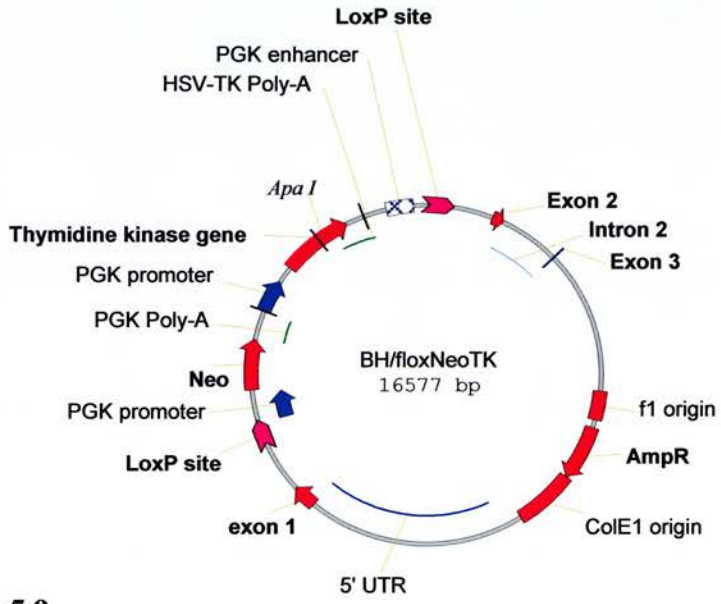


Figure 5.9

Inserting the floxed Neo/TK cassette into the 8kb fragment

This figure shows where the blunted BHI/Xho fragment of the floxed Neo/TK cassette was cloned into a blunted BlnI site in Intron I of the BHI 8kb λ clone in pBS backbone.

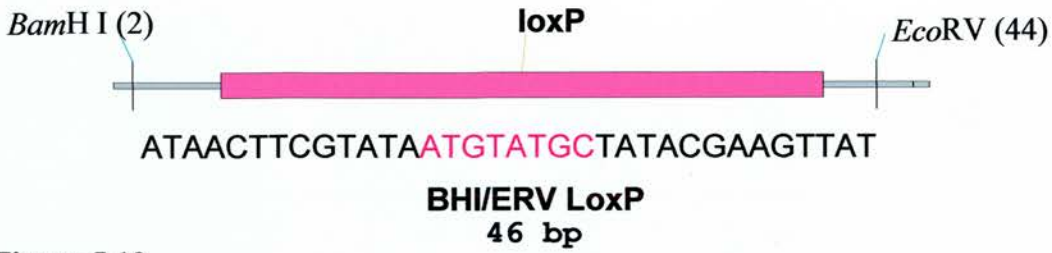


Figure 5.10

loxP sites

This diagram shows a single loxP site with EcoRV and BamHI sites on either end. These sites aid in cloning the loxP sites into the targeting vector and also help determine orientation of the loxP site.

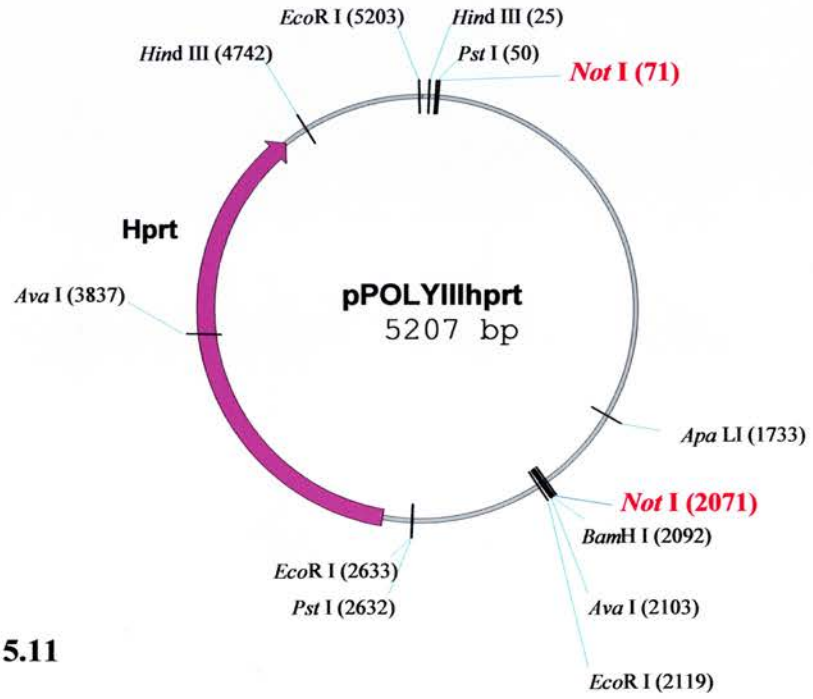


Figure 5.11

pPOLYIII - the HPRT cassette

This figure shows the pPOLYIII-HPRT cassette before being excised from the backbone and cloned into the Smad4 pTarVec targeting plasmid. Note the NotI sites in bold which were used to liberate the Hprt gene from the pPOLYIII backbone.

recombination) then the HPRT cassette would remain present in the DNA and could be selected against using the drug 6-thioguanine (6-TG). 6-TG is a guanine analogue and substrate for the *Hprt* enzyme it undergoes a phosphoribosylation reaction to generate 6-thioguanine monophosphate (6-thioGMP). High levels of 6-thioGMP within the cells inhibit the biosynthesis of guanine nucleotides and this inhibition eventually results in the death of all cells that express *Hprt* (Miech *et al,* 1967) If homologous recombination occurs then the HPRT cassette will be lost and selection against the presence of it will result in a higher proportion of correctly targeted clones. This strategy of counter-selection works only with embryonic stem cells which are already HPRT-deficient as HM-1 ES cells are. The HPRT cassette can be excised from the pPOLYIII/*Hprt* vector (a kind gift from Professor David Melton) by NotI restriction digest (**Figure 5.11**) and this was cloned into the NotI site of the modified 8kb/pBS plasmid to create the final targeting vector, pTarVec (**Figure 5.12**). The overall strategy to successfully introduce loxP sites on either side of exon 1 of the Smad4 locus is shown in **Figure 5.13**.

5.2.3 Generation of correctly targeted embryonic stem (ES) cells

The pTarVec plasmid was transfected into bacteria such that 150μg of ClaI linearised plasmid DNA was ready for electroporation into ES cells. The cells were electroporated as described in **Chapter 2** and then plated. 24 hours after electroporation the medium was removed and G418 selection medium added. Two weeks later the colonies were picked, expanded and harvested. DNA was prepared for Southern blot analysis.

5.3 Teviot Place Analyses

5.3.1 Southern Blot analysis of first round targeting

A Southern blot strategy was designed to detect correct integration of the targeting vector at the 3' end of the region of homology. The ES cell DNA generated from clones surviving selection in G418 antibiotic was digested with the restriction

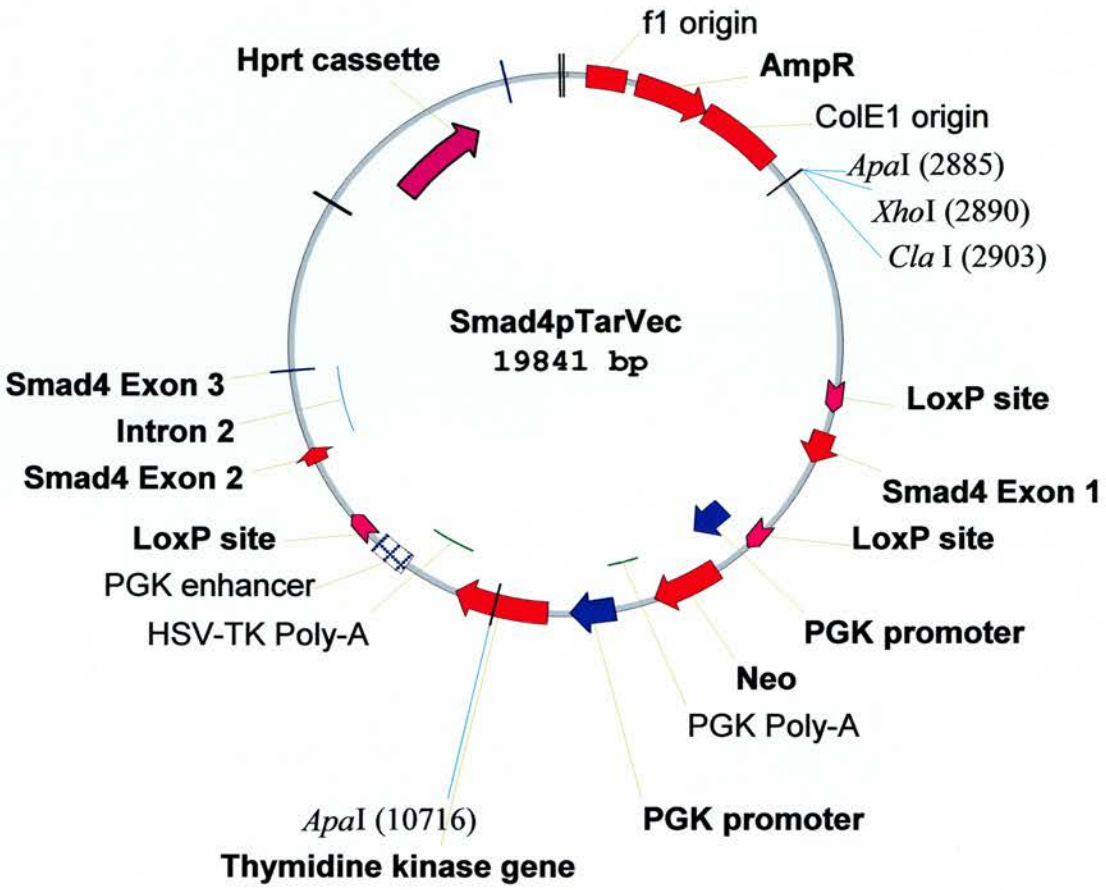


Figure 5.12

pTarVec - the floxed Smad4 targeting vector

This figure shows the complete targeting vector within pBluescript. The targeting vector consists of two loxP sites on either side of exon I of the Smad4 locus and a floxed Neo/TK cassette downstream of exon I. There is an HPRT cassette situated downstream of Smad4 exon 3 (HPRT exons 1-9).

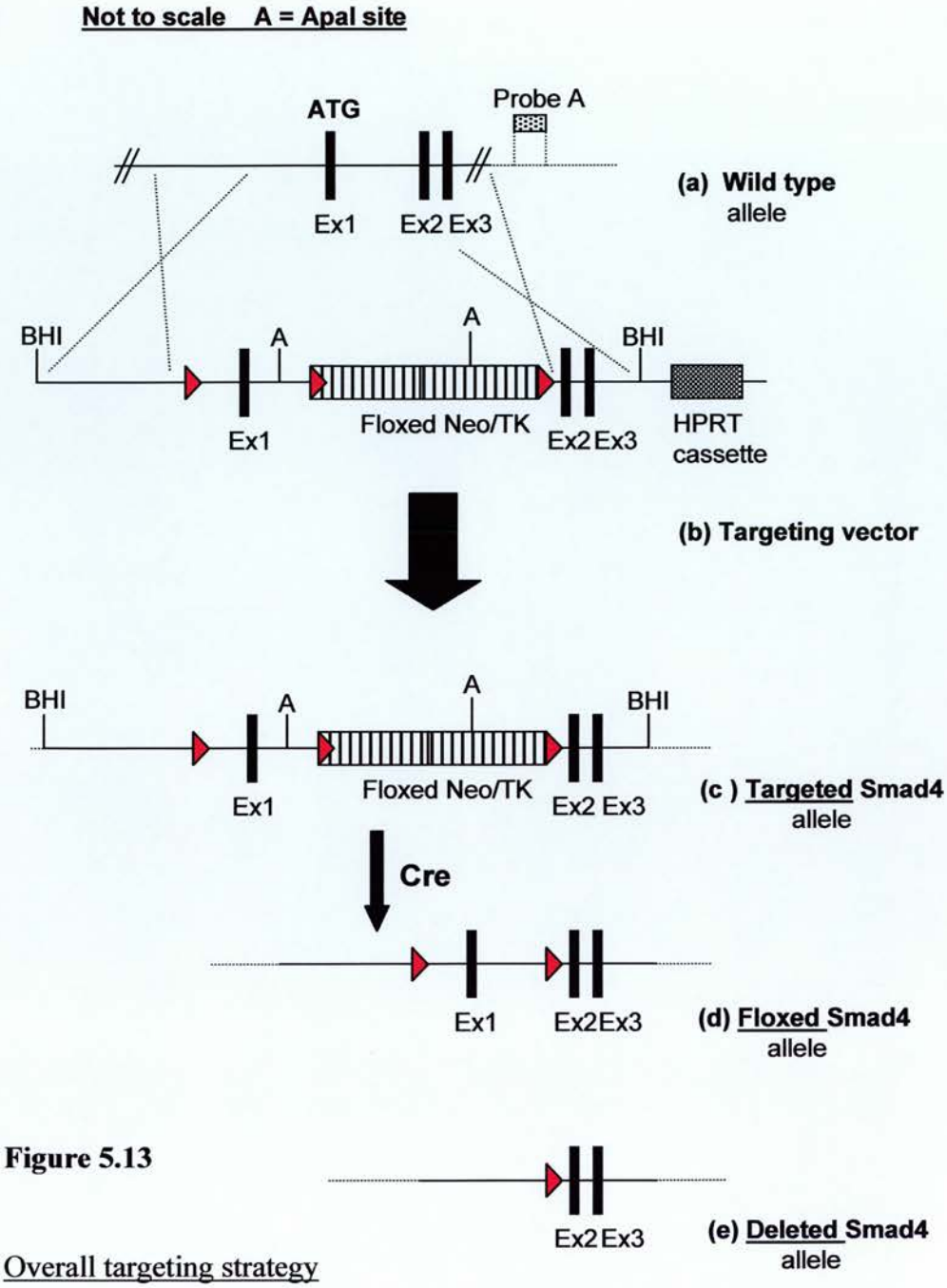


Figure 5.13

Overall targeting strategy

This figure shows the overall strategy employed to achieve conditional deletion of the Smad4 allele. Homologous recombination between the regions of DNA both upstream of the 5' loxP site and downstream of the floxed selection cassette is required to integrate all three loxP sites into the Smad4 locus. Embryonic stem cells with the structure shown in (**d**) will be injected into blastocysts to generate mice bearing a "floxed" Smad4 allele. Crossing of these mice to Cre-expressing transgenic animals will result in deletion of the Smad4 ATG of the floxed allele, essentially making the locus null for Smad4.

endonuclease ApaI and probed with a 2kb BamHI/MunI external probe obtained from restriction digest of the original 15kb λ clone (see **Figure 5.5 (d)** for details). The probe had been shown to bind to a 12kb ApaI fragment generated from wild type DNA and the correct integration of the floxed Neo/TK cassette would introduce a number of new ApaI sites which would shorten the ApaI fragment to approximately 10kb. Southern blot analysis in this way revealed that a number of ES cell clones appeared to have both a wild type ApaI fragment of 12kb and also a targeted allele fragment of 10kb. 50 clones out of 200 hundred picked were analysed in this way and of those approximately 40% contained the targeted allele (**Figure 5.14**).

A Southern blot strategy for the 5' region proved to be more difficult to design and to date this region of the targeted allele has not been confirmed by Southern blot analysis.

5.3.2 Administration of Cre recombinase to remove the floxed Neo/TK

Two approaches to the administration of Cre recombinase were taken. Both utilised the PGK-Cre plasmid created by Dr Tom Gardner at the University of Edinburgh. This plasmid was built from PMC-Cre by the addition of a PGK promoter-Hygromycin resistance gene-PGK enhancer cassette from PHA57 (**Figure 5.15**). The first approach (*Teviot Place Protocol*) was to transiently transfect the correctly targeted first round ES cells with the plasmid using a lipid-based transfection agent. Two clones from the first round of targeting, clone 1 and clone 21 were seeded at low densities in 6 well plates (400 cells per plate and 4000 cells per plate) and transfected with 1μg of PGK-Cre plasmid using the Promega TfX-50 kit. The cells were left in complete medium for one week and only then was medium containing 2μM ganciclovir added. 10 days later colonies were visible to the eye and were marked and picked as described before. DNA was made from the ES cells for PCR analysis.

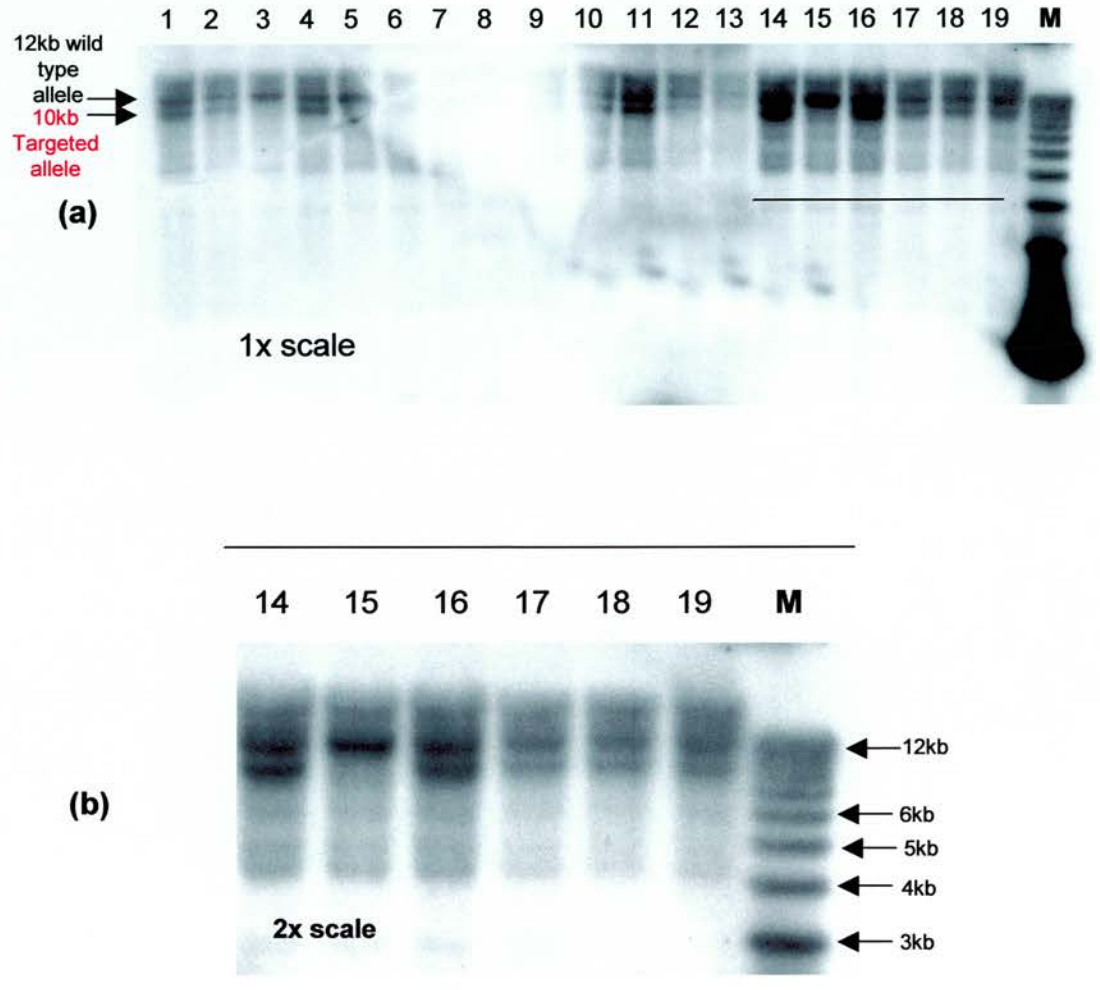


Figure 5.14

Apa I digest Southern blot probed with 1.8kb BHI/MunI 3' probe

Figure (**a**) shows the first 19 samples of a Southern blot to detect correctly targeted floxed Smad4 clones. The ES cell DNA was digested with ApaI and probed with an external 3' probe created from a restriction digest with BHI and MunI. ApaI restriction digest should give a 12kb wild type band and 10kb targeted band via the introduction of a new ApaI site within the Neo/TK. From the clones tested approximately 40% were correctly targeted by this criterion. Figure (**b**) shows an enlargement of a section of the blot clearly showing the two bands at 12kb and 10kb.

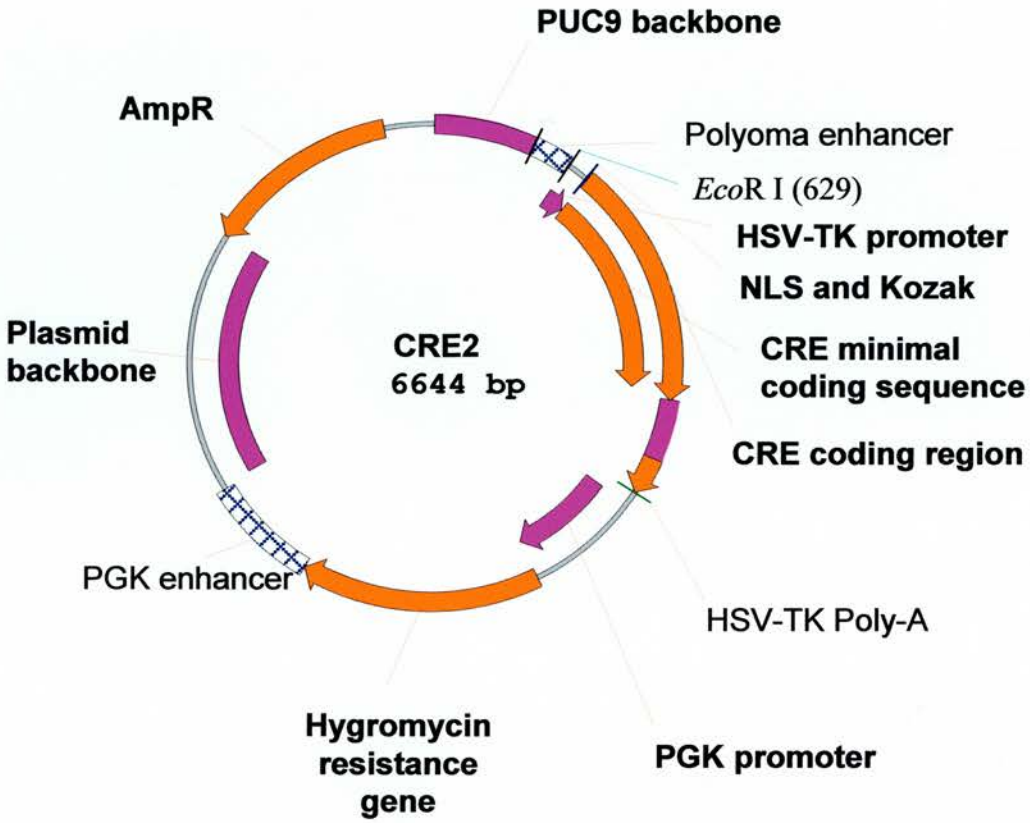


Figure 5.15

The PGK-Cre plasmid

This plasmid was built from PMC-Cre by the addition of a PGK promoter-Hygromycin resistance gene-PGK enhancer cassette from PHA57. The end result can be used in the *in vitro* Cre steps involved to excise the floxed Neo/TK from the targeted allele via recombination in the presence of Cre.

The second approach (*NPU protocol*) was electroporation of the ES cells with the circular PGK-Cre plasmid. The cells and the DNA were mixed together and transferred to an electroporation cuvette where they were electroporated using two pulses at 230V and 500 μFd.

The cells were left to rest at room temperature for 15 minutes before adding complete medium containing LIF. The cells were plated out at 10^4 cells per 10 cm plate and 6 days later LIF complete medium containing 2μM ganciclovir was added and the medium was changed as required. By day 15 colonies were visible so 50 colonies were selected from the plates and screened for the presence of the targeted allele by PCR.

5.3.3 PCR Analysis of Cre-mediated deletion of the floxed Neo/TK cassette

Two PCR analyses were carried out to detect a targeted clone which could be used for injection into blastocysts. The first was to detect the presence of the 5' single loxP site upstream of Exon I. This PCR is internal to the targeting vector but is still important to confirm that there are two bands present to ensure that recombination had occurred upstream of this loxP site and that this loxP site had therefore been incorporated into the target locus; one wild type band and one targeted band. This PCR utilised two primers designed from sequence data from either side of the single loxP site. The primers "loxP up" and "380rev (2)" are designed to amplify a band of approximately 220bp with wild type DNA and 280bp when a single loxP site is present (**Figure 5.16**). The second PCR was designed to test that the floxed Neo/TK had been excised properly after addition of the Cre recombinase. This PCR utilised two primers designed from known sequence on either side of the cassette. "DPC4 Intron" and "DPC42" are designed to amplify a wild type band of approximately 230bp whereas the presence of a single loxP site remaining after recombination between the two loxP sites at either side of the cassette should increase this band by approximately 70bp to 300bp (**Figure 5.17**). No signal can be generated from the Neo/TK-containing allele before recombination as this would involve PCR over a

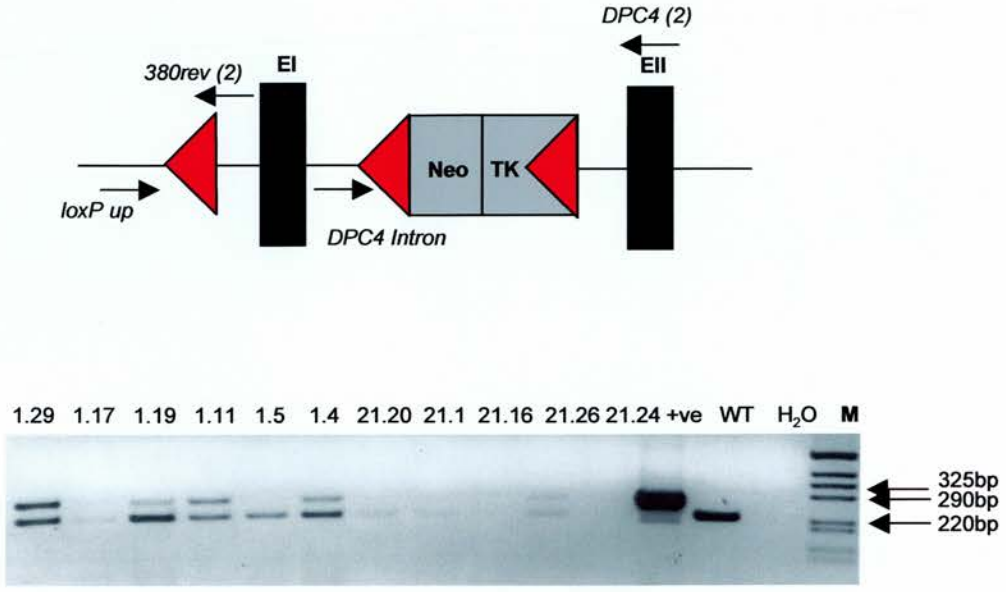


Figure 5.16

PCR with loxP up and 380rev (2) primers

This figure shows the results of a PCR to detect the 5' loxP site within the targeted allele. ES cell clone 1 and 21 (both positive by Southern Blot analysis) were tested after exposure to Ganciclovir to remove the floxed Neo/TK. Clones correctly targeted at the 5' end show two bands, one WT band of approximately 220bp and a loxP site band of approximately 280bp.

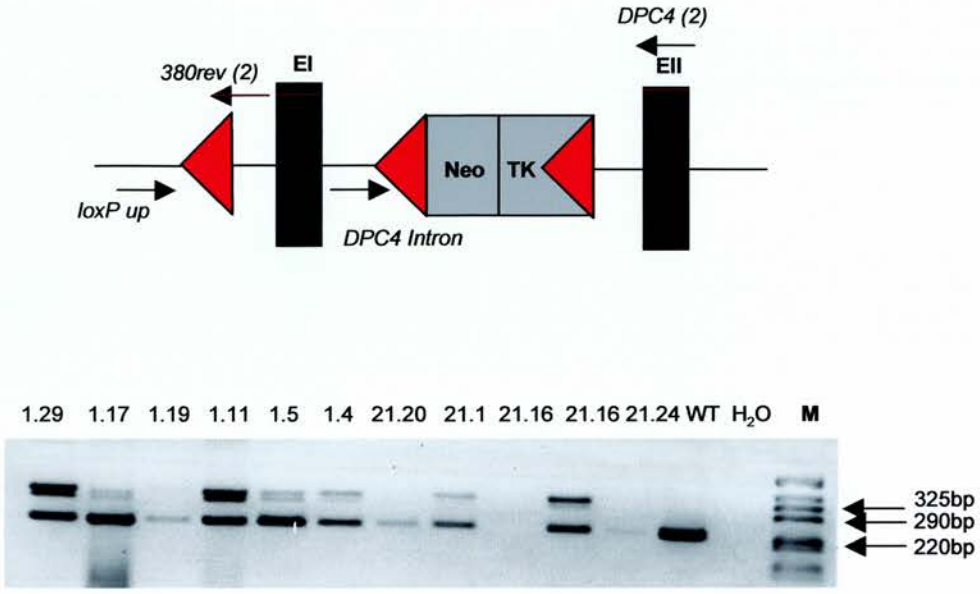


Figure 5.17

PCR with DPC4 Intron and DPC4 (2)

This figure shows the results of a PCR to identify correctly targeted clones which have lost the floxed Neo/TK cassette as a result of exposure to Cre recombinase and selection in ganciclovir. Positive clones will show two bands; one WT band of approximately 230bp and the targeted allele band of approximately 300bp. The doublet in the upper band is thought to be due to loxP secondary structure effects during the PCR reaction.

5kb region of DNA, which is unlikely with the short extension times used in this PCR protocol.

The results of the PCR analyses on Teviot Place-derived ES cell clones were that clone 1.29 was assumed to be correctly targeted with single loxP sites remaining which flank exon I on one allele of the murine Smad4 locus. Clone 1.29 was subsequently assessed for karotype.

5.3.4 Karyotyping correctly targeted ES cell clones

Standard methodology was used to karotype ES cell clones (See **Chapter 2**). Preparation of metaphase spreads was achieved by fixing the cells and dropping them on to a microscope slide from a height of about 30 cm. After staining the slides with Giemsa it was possible to count chromosomes under the microscope.

Clone 1.29 had a modal chromosome number of 40, with approximately 65% of the spreads counted having the modal number. 50 fields of cells were counted. This figure gives an indication of the ability of the ES cell line to be germline competent. Although only 65% of the cells counted have a normal modal chromosome number it has been previously shown that HM-1 ES cells with karyotypes of this nature can indeed produce good chimeras and can generate germline transmission of the targeted allele (*Manson J., personal communication*).

Clone 1.29 was subsequently grown to yield a number of large flasks of cells for freezing purposes. At this point the passage number of the clone was 23.

Mrs Jennifer Doig and Ms Tracey Higgins carried out blastocyst injections at the department of Pathology, Teviot Place. A vial of ES cells was defrosted 2 days prior to injection and grown in a 25cm^2 flask in medium supplemented with LIF. The ES cells were harvested as usual on the day of injection and injected into blastocysts obtained from superovulated C57BL/6 pseudopregnant female mice. 15-30 embryonic stem cells were injected into each blastocyst and on average 10-15

blastocysts were returned. From 6 injection sessions only one chimera was produced with more than 25% chimerism (**Figure 5.18**). This one male was allowed to mate repeatedly, but failed to produce any germline transmission of ES cell derived coat colour markers. After this male had fathered 50 negative pups this chimera was considered not to transmit ES cell coat markers and the situation re-assessed.

It was decided to re-target ES cells at the Neuropathogenesis Unit (NPU), Kings Buildings, Edinburgh for a number of reasons; firstly non-targeted test ES cells at Teviot Place were failing to colonise the germline even when excellent chimeras had been produced, suggesting that there might be a problem with the ES cells, the culture technique or the pseudopregnant females' blastocysts. Secondly, the NPU had been consistently producing high quality chimeras (90%+) and subsequent germline transmission using different protocols and methods even although their ES cells were derived from the same stock as those being used at Teviot Place.

5.4 NPU Analyses

5.4.1 Targeting the Smad4 locus at the NPU

Targeting the Smad4 locus with the pTarVec floxed targeting vector commenced at the NPU in late 1999/early 2000. I have outlined some of the differences in the protocols used in the Materials and Methods section of this thesis (**Chapter 2**) but essentially the main differences were the medium used, the electroporation protocols and the *in vitro* Cre step (electroporation instead of transient transfection). However, the same re-derived HM-1 ES cell line was used and the end result of the experiment was effectively the same in terms of the number of correctly targeted clones at each stage of the experiment. Re-derived HM-1 ES cells were electroporated with 150 μg of linearised pTarVec plasmid as described in **Chapter 2**. Unfortunately a number of dishes were destroyed by a fungal infection but of only ten ES cell clones picked 4 showed the presence of the targeted allele as described in section **5.3.1** of this chapter. These targeted clones were then subjected to electroporation with the PGK-Cre plasmid and from these a number of clones which survived selection in

Figure 5.18

Teviot Place Chimera No. 1

This photograph shows a chimera generated from correctly targeted floxed Smad4 ES cells. The animal is chimeric in coat colour only on the ventral side and never produced germ-line offspring.

ganciclovir were tested for the presence of the floxed Smad4 genomic structure using the PCR described in section **5.3.3**. One clone, numbered 49.8, was found to contain the two loxP sites indicative of a floxed allele.

However, two differences were that the final clone, Clone 49.8, had a karyotype of 40 in 72% of cells counted (50 fields of cells counted) and the clone was frozen down for blastocyst injections at a lower passage number – passage 19.

Dr Joe Mee of the Centre for Genome Research, Kings Buildings, Edinburgh (CGR) carried out the ES cell injections into blastocysts from naturally mated C57BL/6 female mice and transferred them into pseudopregnant F1 recipient female mice. From the first injection session 3 chimeras were produced, two of which were termed "highly chimeric" on the basis of coat colour and one mouse demonstrated medium to low chimerism. From a second injection session another 3 chimeras were produced (**Figure 5.19**). The mice from the first injection session were then mated to MF1 mice to identify germline transmission of the ES cell markers. Pups born from these matings showed germline transmission of the ES cell markers in 50% of two of the three litters and 100% of the third litter (**Figure 5.20 (a)**). Seven pups from a total of 13 were tail-tipped and PCR analysis for the presence of the two loxP sites on either side of exon I carried out from the tail tip DNA (**Figure 5.20 (b) and (c)**).

The results demonstrate that of the seven pups genotyped in this way, mouse **number 1** shows germline transmission of the floxed Smad4 allele, the pattern of bands observed in the PCR being identical to those of the ES cells used for targeting, clone 49.8.

5.5 Discussion

5.5.1 Floxed Smad4 mice – where next?

The generation of mice with a floxed Smad4 allele has been a major part of this project and although this initial stage of this work is now complete, this marks the

Figure 5.19

CGR floxed Smad4 chimeras

Targeted floxed Smad4 ES cells were injected into C57Bl blastocysts by Dr Joe Mee at the Centre for Genome Research (CGR) at Edinburgh University. The blastocysts were placed into the uterine horns of 2.5dpc pseudopregnant F1 mice and pups born approximately 18 days later. Three chimeras were produced from the first injections, one chimera was classed as low to medium levels of chimerism (pictures **1a,1b**) and two of which were classed as having high levels of chimerism (pictures **2a-3b**). A second round of injections produced a number of highly chimeric pups (picture **4**).

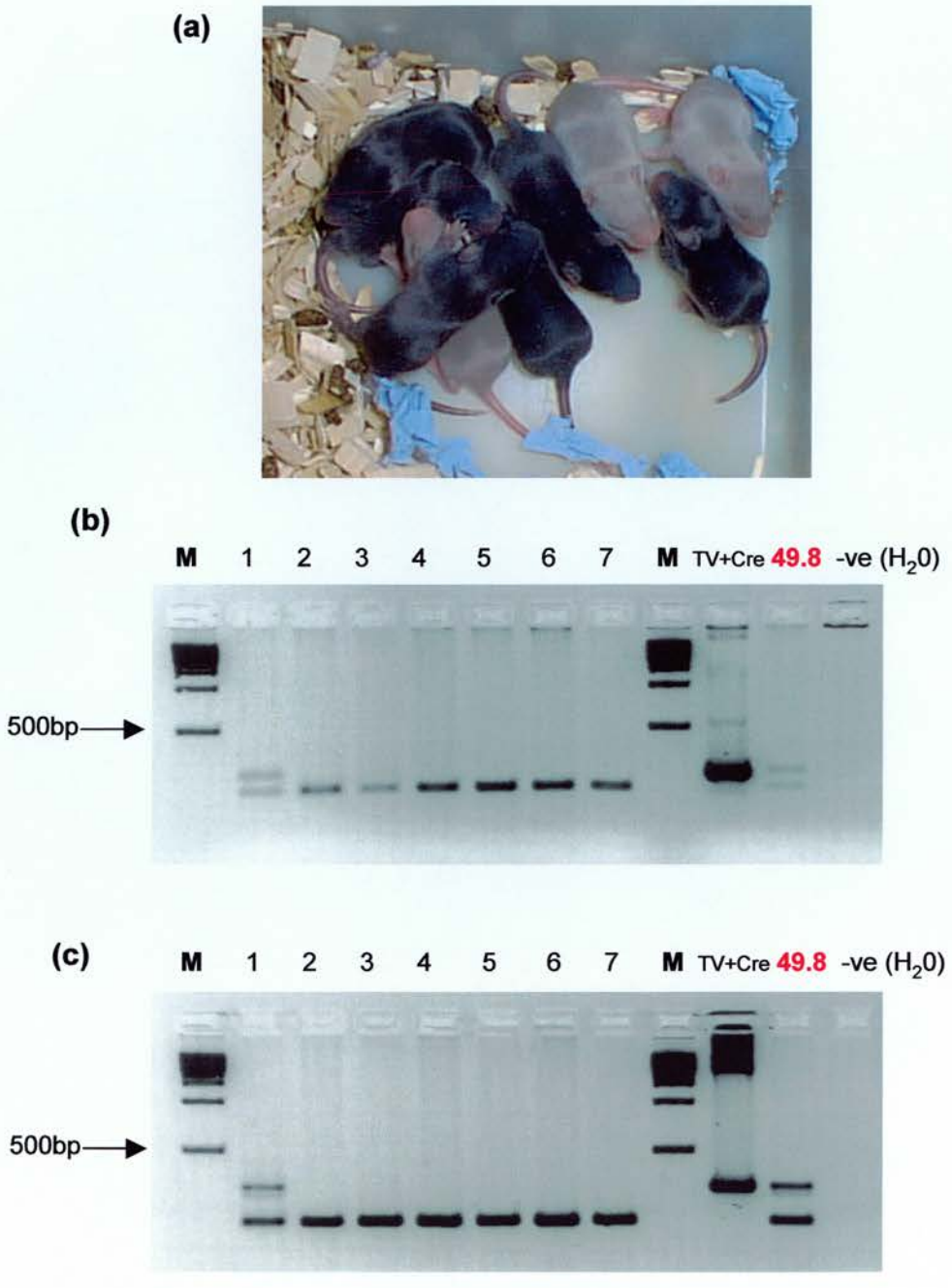


Figure 5.20

Germline Transmission of the Floxed Smad4 allele

This figure shows **(a)** pups born from a floxed Smad4 male chimera x MF1female, which resulted in three pups showing germline transmission (grey pups). Figure **(b)** shows a PCR reaction to determine the presence of the single loxP site upstream of exon I and Figure **(c)** shows a PCR reaction to determine the single loxP site downstream of exon I.. Mouse **no. 1** shows the two bands identical to those observed in the ES cell clone 49.8

beginning of the biological analysis. Germline transmission of the targeted Smad4 allele represents the first stage of a breeding programme required to create a null Smad4 allele only in the mammary epithelium of the mouse. The next steps are:

1) Mating the $Smad4^{+/fl}$ progeny of this cross to $Smad4^{+/-}$ animals.
2) Genotyping the progeny from this cross to determine which animals are $Smad4^{fl/-}$
3) Mating $Smad4^{fl/-}$ animals to BLG-Cre transgenic mice.
4) Analysing the progeny of this cross which should have deleted the floxed allele and effectively be $Smad^{-/-}$ in the mammary gland only.

Analysis of the mammary gland-specific null Smad4 animals will probably take on a similar pattern to analysis of the BLG-Smad4 transgenic animals. This will involve wholemount and H&E analysis of both wild type and Smad4 null mammary glands and an extensive biochemical analysis of genes both related to and downstream of the TGF-β signalling pathway. One of the first questions to be addressed will be what levels of residual Smad4 are observed in the gland after exposure to Cre. Previous studies have shown that the BLG-Cre transgene reproducibly delivers 80% deletion of floxed genes (Chapman *et al*, 1999) and a similar level of recombination is predicted for Smad4. Western blotting with a Smad4 antibody will be used to determine exactly how much Smad4 protein remains in the mammary gland and where it is located.

Predicting a phenotype for this experiment is difficult given that there are currently no other mammary gland-specific floxed TGF-β-related genes in the literature, and the majority of conventional knockouts of TGF-β genes result in embryonic lethality. However, given the widespread role of TGF-β in control of epithelial cell growth the loss of Smad4 (and possible loss of some TGF-β signalling pathway components) is predicted to lead to dysregulation of the mammary gland. How this dysregulation presents itself is more difficult to predict as we currently have little information on what role Smad4 actually plays in the mammary gland. However, the immediate

hypothesis to be tested, given the observed up-regulation of Smad4 at the start of involution (See **Chapter 6**) and the preliminary observation that over-expression of Smad4 may lead to accelerated involution (See **Chapter 4**), is that loss of Smad4 will impair the normal programme of involution.

The availability of mice bearing a conditional Smad4 allele will also allow an assessment of the potential of Smad4 to act as a tumour suppressor within any given tissue, dependent upon the pattern of Cre mediated excision. Loss of Smad4 has been reported to occur in a range of tumour types (Xu *et al*, 2000 and Schutte *et al*, 1996, Duff and Clarke, 1998), and it is clear that Smad4 is a potent tumour suppressor in the murine intestine (Takagi *et al*, 1996, Takaku *et al*, 1999 and Xu *et al*, 2000). The availability of mice bearing a conditional allele will allow these roles to be dissected further *in vivo*. In particular, this strain will permit an analysis of tumour suppressive activity within the mammary epithelium. Schutte *et al* (1996) reported that Smad4 was lost in 12% of human breast cancers, suggesting that Smad4 may be important in preventing a subset of neoplasias in the breast. This data was, however, derived from a limited analysis of tumour cell lines, and it is questionable whether this truly reflects a role for Smad4 in suppressing neoplasias. Conditional inactivation of Smad4 in the mammary epithelium will directly address this question and may also generate a murine model for this subset of human breast cancers.

Numerous publications have demonstrated that loss of responsiveness to TGF-β is a key factor in the progression to neoplasia (for review see Gold, 1999), and that this may be mediated through both altered cellular proliferation and cell death. It is also clear that functional Smad4 is an essential component of the normal TGF-β response. The availability of mice bearing a conditional allele will allow a direct analysis of the mechanisms through which perturbation of the TGF-β response can initiate and promote neoplasia, and a precise assessment of the role of Smad4. The development of these strains will not only permit a fundamental understanding of the significance of signalling pathways in human cancer, but may also provide valuable models of specific human cancers, such as of those of the breast and intestine.

Chapter 6 – Smad-STAT Interactions in the Mammary Gland

6.1 Introduction

6.1.1 Mammary Gland Involution

As described in detail in **Chapter 1** mammary gland involution proceeds through two distinct stages and both these stages exhibit characteristic changes in gene expression or activity. First stage changes include up-regulated expression of sulphated glycoprotein-2 (SGP-2), tissue inhibitor of metalloproteinases-1 (TIMP), interleukin -1ß converting enzyme (ICE), and cell cycle control proteins (c-Jun, junB, junD, c-fos and c-myc) and decreased expression levels of milk protein genes. Second stage changes include increased expression levels of matrix metalloproteinases gelatinase A and stromelysin-1 and serine protease urokinase-type Plasminogen activator (**Figure 6.1**)

Programmed Cell Death (PCD) of individual alveolar cells during the first days of involution is correlated with increased expression levels of the death inducers bax and bcl xshort (bcl-xS) as compared to bcl-xlong (bcl-xL), both of which are members of the bcl-2 family. Importantly, this phase of PCD is p53 independent. Changes in activity of two STAT family members accompany mammary gland involution: decreased activity of the prolactin signalling molecules STAT5a and STAT5b and activation of STAT3.

Recently the Smad gene family have been identified as mediators of TGF-β superfamily signalling and have been implicated in mediating epithelial cell apoptosis (Brodin *et al*, 1999, Yanagisawa *et al*, 1998 and Ishisaki *et al*, 1998). This chapter aims to examine the *in vivo* potential for such a role during mammary gland involution and to elucidate interactions with other signalling pathways.

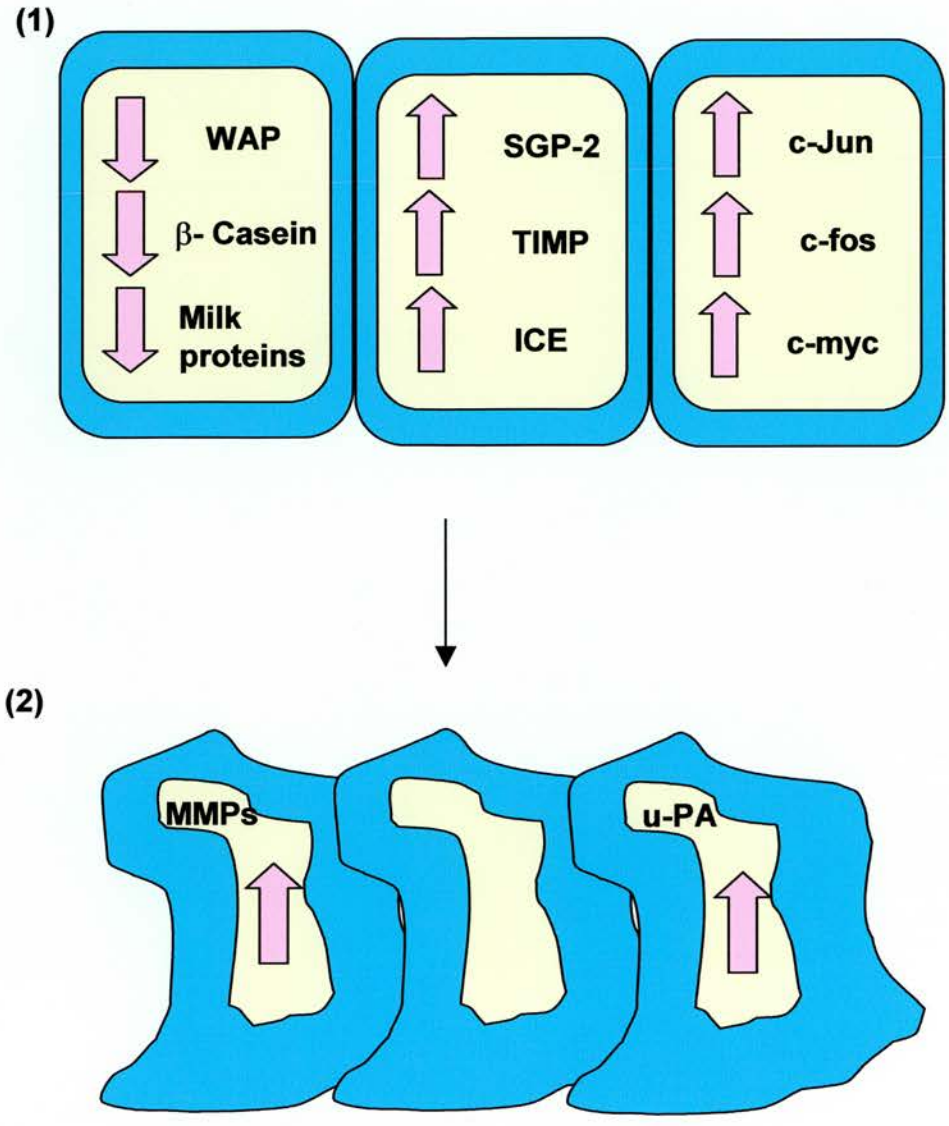


Figure 6.1

The two stages of mammary gland involution

(**1**) In the first stage of involution alveolar cells undergo programmed cell death, but there is no remodelling of the lobular-alveolar structure. First stage changes include up-regulated expression of inhibitors of proteases, apoptosis-associated genes and cell cycle control proteins and this is concomitant with decreased expression levels of various milk protein genes (**2**) During the second stage of involution the lobular-alveolar structure of the gland is obliterated as proteinases degrade basement membrane and extra-cellular matrix (ECM).

6.1.2 Smads and STATs and Apoptosis

Smads are the family of signalling intermediates involved in signal transduction from Transforming Growth Factor β (TGF-β) superfamily members (Massague 1998; Heldin *et al*, 1997; Christian and Nakayama, 1999) and, as described in **Chapter 1** of this thesis, this group of proteins was discovered in 1996 when the first member of the group, *Mothers against decapentaplegic* (*Mad*) was identified during a screen for mutations which enhanced *dpp* mutants (Hahn *et al*, 1996). Smads have been shown to play a multitude of important biological roles including the ability to function as tumour suppressors (Duff and Clarke, 1998; Riggens *et al*, 1997). This ability may arise as a consequence of having been shown to initiate a number of cellular responses including apoptosis. This is perhaps not surprising, as members of the TGF-β superfamily have long been implicated in the induction of cellular apoptosis through activation of proteins associated with cell cycle arrest such as p21WAF1/Cip1 and p27/Kip1 (Rosfjord and Dickson, 1999). Individual Smads have also been shown to initiate apoptosis *in vitro*, thought to be mediated through an up-regulation of p21 (Dai *et al*, 1999; Grau *et al*, 1997 and Chiao *et al*, 1999) and are also up regulated during involution in the prostate (Brodin *et al*, 1999). Recently Smads have been shown to interact with another family of intermediate signalling molecules, STATs (**S**ignal **T**ransducer and **A**ctivator of **T**ranscription). STATs are a family of latent transcription factors which are activated in response to many cytokines and growth factors (Schindler, 1998) and have been shown to be capable of both activating and suppressing gene transcription (Fukada *et al*, 1998). Members of the STAT family of transcription factors are present in species as diverse as mammals, insects and slime moulds. Discovered as mediators of interferon-induced signals, the STATs were later shown to drive many different ligand-induced responses through receptor-induced tyrosine phosphorylation and dimerisation. STAT1 also functions as a transcription factor, essential for the efficient constitutive expression of certain genes, without needing tyrosine phosphorylation, and phosphorylated STAT1 dimers mediate suppression – rather than activation – of some genes. STATs are present in the cytoplasm of untreated cells in multiprotein

complexes, which might aid in their nuclear translocation and differential binding to DNA, thus contributing to the specificity of STAT action. STATs are activated in response to many different cytokines and growth factors, and they can function as either growth suppressors (Kaplan *et al,* 1998) or growth promoters (Catlett-Falcone *et al*, 1999) depending on the particular STAT and the promoter sequence to which it binds. STATs regulate cellular functions as diverse as embryonic development (Nemetz and Hocke, 1998) cell death (Kumar *et al*, 1997) and cell migration (Mohanty *et al*, 1999) and to date at least seven mammalian STATs have been identified, ranging in size from 90 to 115 kDa. The STAT signalling pathway has been well defined as an archetypal membrane to nucleus signal transduction pathway which utilises the Janus kinase (JAK) proteins to achieve a cascade of phosphorylation signals leading to gene transcription. The classic activation pathway involves the cytokine ligand IFNγ which induces oligomerisation of the IFNγ receptor subunits, leading to phosphorylation and activation of the preassociated JAKs, which then phosphorylate the IFNγ receptor, allowing STAT1 to dock. STAT1 is then phosphorylated and forms a homodimer that migrates to the nucleus (**Figure 6.2**). TGF-β and Smads have been shown to be inhibited by the IFN-γ/STAT pathway (Ulloa *et al*, 1999). Here IFN-γ inhibits the TGF β-induced phosphorylation of Smad3 and its attendant events, namely, the association of Smad3 with Smad4, the accumulation of Smad3 in the nucleus, and the activation of TGF-β-responsive genes. Acting through JAK1 and STAT1, IFN-γ induces the expression of Smad7, an antagonistic Smad, which prevents the interaction of Smad3 with the TGF-β receptor. These observations were the first to indicate a possible mechanism of transmodulation between the STAT and Smad signal-transduction pathways. A second recent paper enforces this concept that Smad and STAT pathways are closely linked with the observation that *in vitro* Smad1 and STAT3 can interact through the transcriptional activator p300 (Nakashima *et al*, 1999).

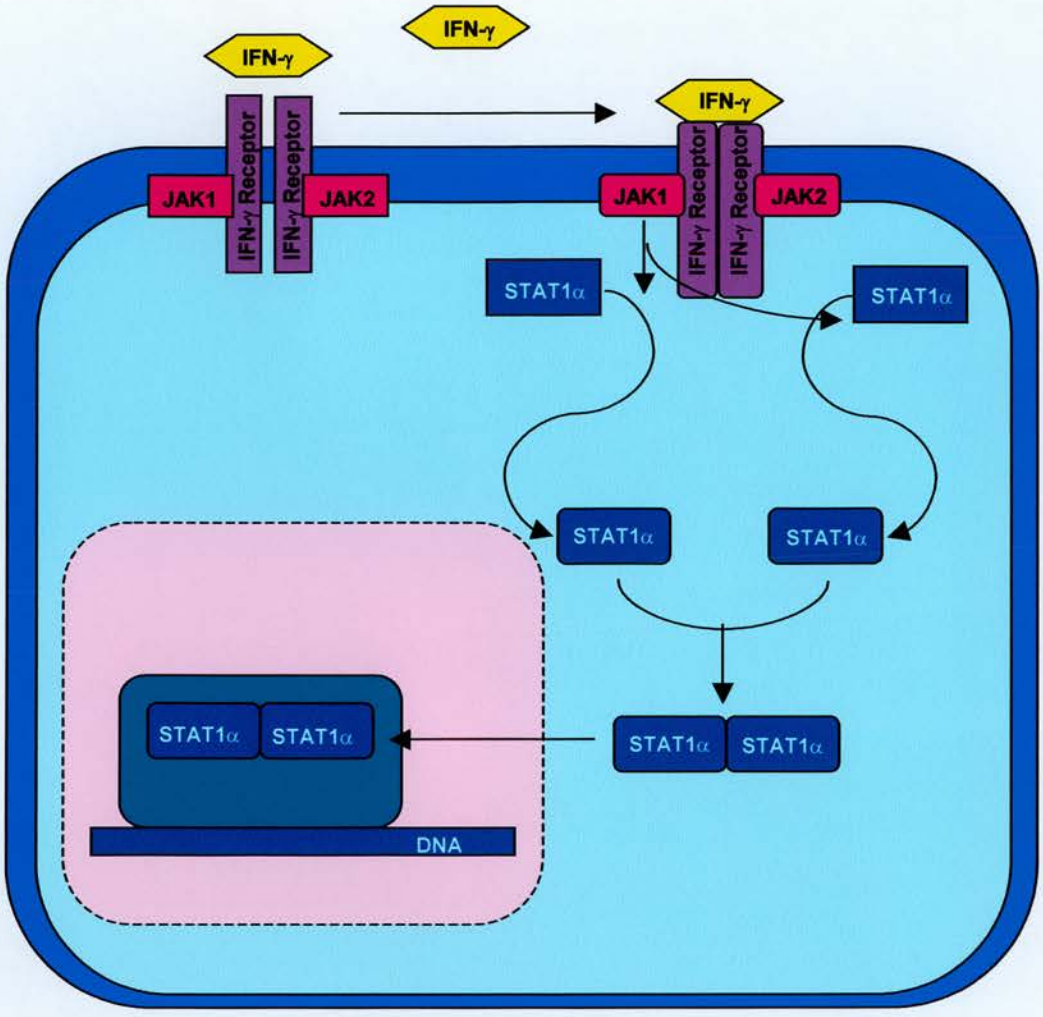


Figure 6.2

STAT signalling pathway

The classic activation pathway involves the cytokine ligand IFNγ which induces oligomerisation of the IFNγ receptor subunits, leading to phosphorylation and activation of the preassociated JAKs, which then phosphorylate the IFNγ receptor, allowing STAT1 to dock. STAT1 is then phosphorylated and forms a homodimer that migrates to the nucleus

6.1.3 Co-activators and Complexes

p300 and CBP are large proteins, which are structurally and functionally conserved in mammals, *C. elegans* and *Drosophila,* and which have been demonstrated to serve coactivator roles for several classes of transcription factors. p300 is a close homologue of CREB- binding protein (CBP), a large protein discovered in 1998 to associate with cAMP response element binding protein (CREB) in response to cAMP signalling (Chrivia *et al*, 1993). p300 was purified as a cellular binding protein of the adenoviral E1A protein (Eckner *et al*, 1994) and these two molecules have been implicated in functions of a large number of regulated transcription factors based primarily on two criteria; physical interaction and the ability to potentiate transcription when overexpressed (Gerritsen *et al*, 1997). One reason for these molecules' importance for so many different transcription factors may be the finding that both CBP and p300 harbour histone acetyl transferase (HAT) activity (Bannister and Kouzarides, 1996 and Ogryzko *et al*, 1996). Compact chromatin structure imposes a repression effect on transcription *in vivo*. Histone acetylation results in decreased affinity between core histone subunits and DNA, facilitating entry of core transcription factors and is thus correlated with increased transcriptional activation.

Smad1, STAT3 and p300 were found to bind in a complex to the GFAP promoter in neural cell types and this was the first indication that Smads and STATs can physically interact (Nakashima *et al*, 1999). p300 was shown to interact physically with STAT3 at its amino terminus in a cytokine stimulation-independent manner, and with Smad1 at its carboxyl terminus in a cytokine stimulation-dependent manner. The formation of a complex between STAT3 and Smad1, bridged by p300, was then elegantly shown to be involved in the cooperative signalling of LIF and BMP2 and the subsequent induction of astrocytes from neural progenitors. A third more recent paper has shown that Smad1 interacts with p300 and that this interaction can synergistically activate gene transcription (Pearson *et al*, 1999). This group demonstrated that Smad1 interacts with p300 and CBP both *in vitro* and *in vivo*. The N- and C-termini of Smad1 negatively interfere with binding to p300/CBP, and the C-terminal half of Smad1 harbours two interaction domains both binding

to the same region near the C-terminus of p300/CBP. They showed that Smad1 can stimulate gene transcription in a cooperative fashion with p300/CBP and that phosphorylation of Smad1 enhances its binding to CBP and thereby further stimulates Smad1-dependent transcription. These results were the first to provide a mechanism of how a phosphorylated Smad might regulate gene activity by interaction with p300/CBP and other factors associated with these bridging coactivators.

STATs have been implicated in regulating apoptosis in several systems, but more recently have been shown to be essential for co-ordinated mammary gland involution where there is extensive apoptosis of the secretory epithelial cells (Chapman *et al*, 1999). This study utilised an animal model of conditional deletion of STAT3 in the mammary epithelium during the time of lactation and involution. Following weaning, a decrease in apoptosis and a dramatic delay of involution occurs in STAT3 null mammary tissue. No marked differences were seen in the regulation of STAT5, Bcl-x(L), or Bax in the absence of STAT3. Functional activation of STAT1 and increases in levels of p53 and p21 occurred and may act as compensatory mechanisms for the eventual initiation of involution observed in STAT3 null mammary glands. This was the first demonstration of the importance of a STAT factor in signalling the initiation of physiological apoptosis *in vivo*.

I therefore chose to examine the regulation of Smads during normal mammary gland involution and the interaction they may have with STATs in this environment during that period. This investigation is based on the hypothesis that since STAT3 is known to be important in mammary gland involution (Chapman *et al*, 1999) and TGF-β is known to be both important in mammary gland biology and a known inducer of apoptosis, and that Smads and STATs have been shown to interact; it seems possible that Smads play a role in involution of the mammary gland and that if Smad-STAT interactions are important in the context of mammary gland biology then disruption of one of these components might affect the other. I chose to analyse the effect of perturbed Smad-STAT interactions by utilising a strain of mice carrying a conditional null STAT3 allele. This mouse model has been described before and

briefly consists of a conditional deletion of the *Stat3* gene using the Cre-lox recombination system in which expression of Cre recombinase is directed specifically to mammary epithelial cells by the promoter of the milk protein gene β-lactoglobulin (BLG) (Selbert *et al.* 1998). These BLG-Cre transgenic mice were crossed with mice containing one null *Stat3* allele and one floxed *Stat3* allele in which the loxP sites were inserted around the tyrosine phosphorylation domain to create a functional knockout of *Stat3* (Takeda *et al.* 1998).

6.2 Results

6.2.1 Smads are up regulated at the start of involution

TGF-β-induced apoptosis of epithelial cell types has been well characterised. In order to investigate the potential role of members of the Smad signal transduction pathway in this process I examined levels of various Smads from day 10 of lactation, at which point pups were weaned to induce involution, through to day 6 of involution by Western blot analysis. See **Figure 6.3** for a description of these stages by H&E staining of mammary gland slides. A subset of the Smad proteins was analysed. These comprised Smad1, as an example R-Smad implicated in BMP signalling, Smad2 as an example R-Smad implicated in TGF-β and activin signalling; the principal Co-Smad, Smad4; and the principal anti-Smad implicated in TGF-β and activin signalling, Smad7. I also analysed levels of the activated R-Smad, phospho-Smad2. Analysis of raw data shows that levels of Smad1, Smad2, phospho-Smad2 (R-Smads) and Smad4 (a Co-Smad) are all up regulated at day 2 of involution compared to day 10 lactation. The highest fold induction is observed for Smads 1 and 2 at day 2 of involution. (**Figure 6.4 (a).**). In contrast to this, levels of Smad7, an anti-Smad, remained relatively unchanged during involution, although a slight reduction was seen at day 2 and day 6 (**Figure 6.4 (b).**). Data was also normalised to keratin 18 levels, a marker for epithelial cells, to account for differences in the epithelial cell content of the glands throughout involution attributable to the fact that the STAT3 null mice involute approximately three days later than their wild type

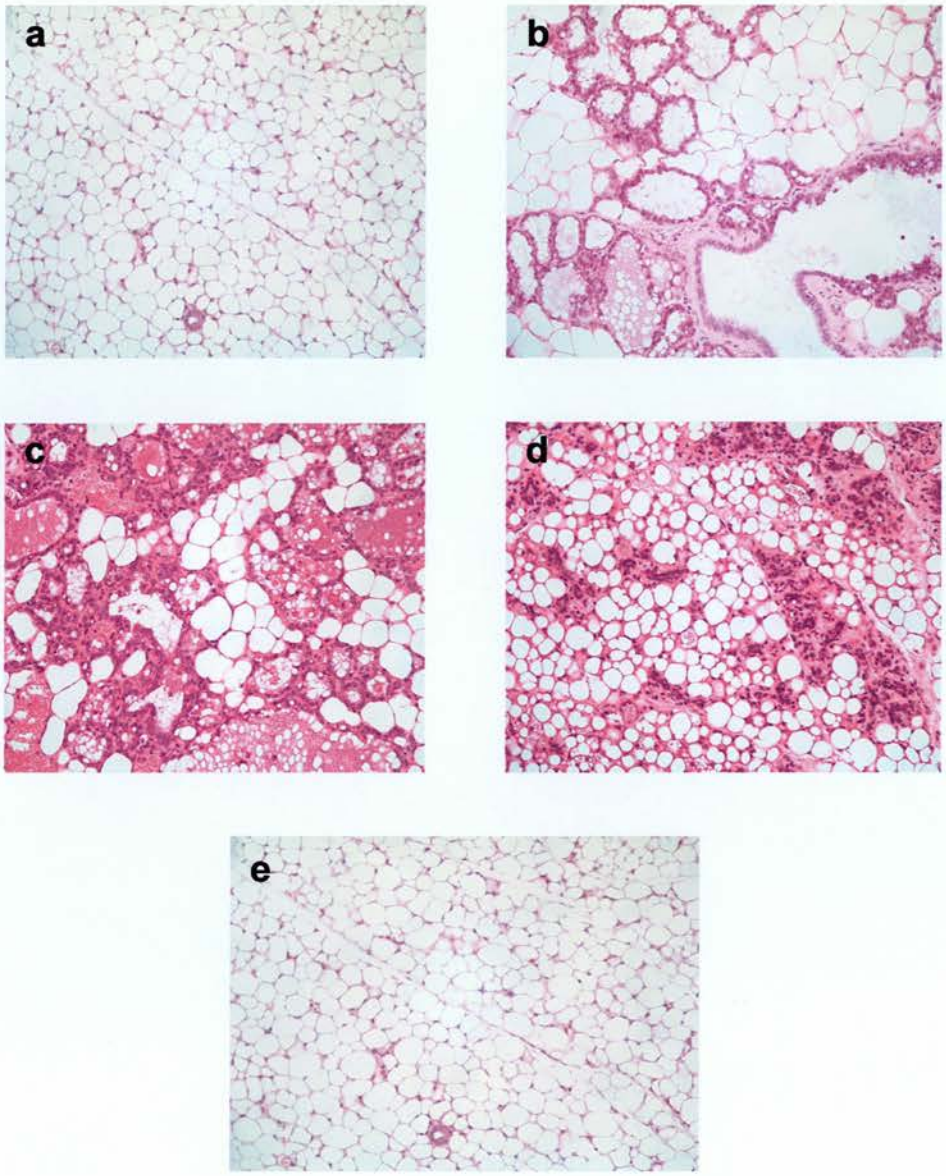


Figure 6.3

H&E stain of mammary gland lactation to involution progression

This figure shows H&E staining of mammary gland in (**a**) virgin mice where the gland is composed of branching epithelium and stromal alveolar structures and (**b**) day 10 lactation where the alveolar structures become filled with milk proteins and the epithelial walls expand. By day 3 (**c**) the lobuloalveolar structure of the gland has started to collapse due to degradation of the extracellular matrix. And by day 6 (**d**), although involution is not complete, the majority of the gland has been remodelled. 10 days after the start of involution (**e**) the gland resembles that of a virgin with the majority of the gland composed of branched epithelium.

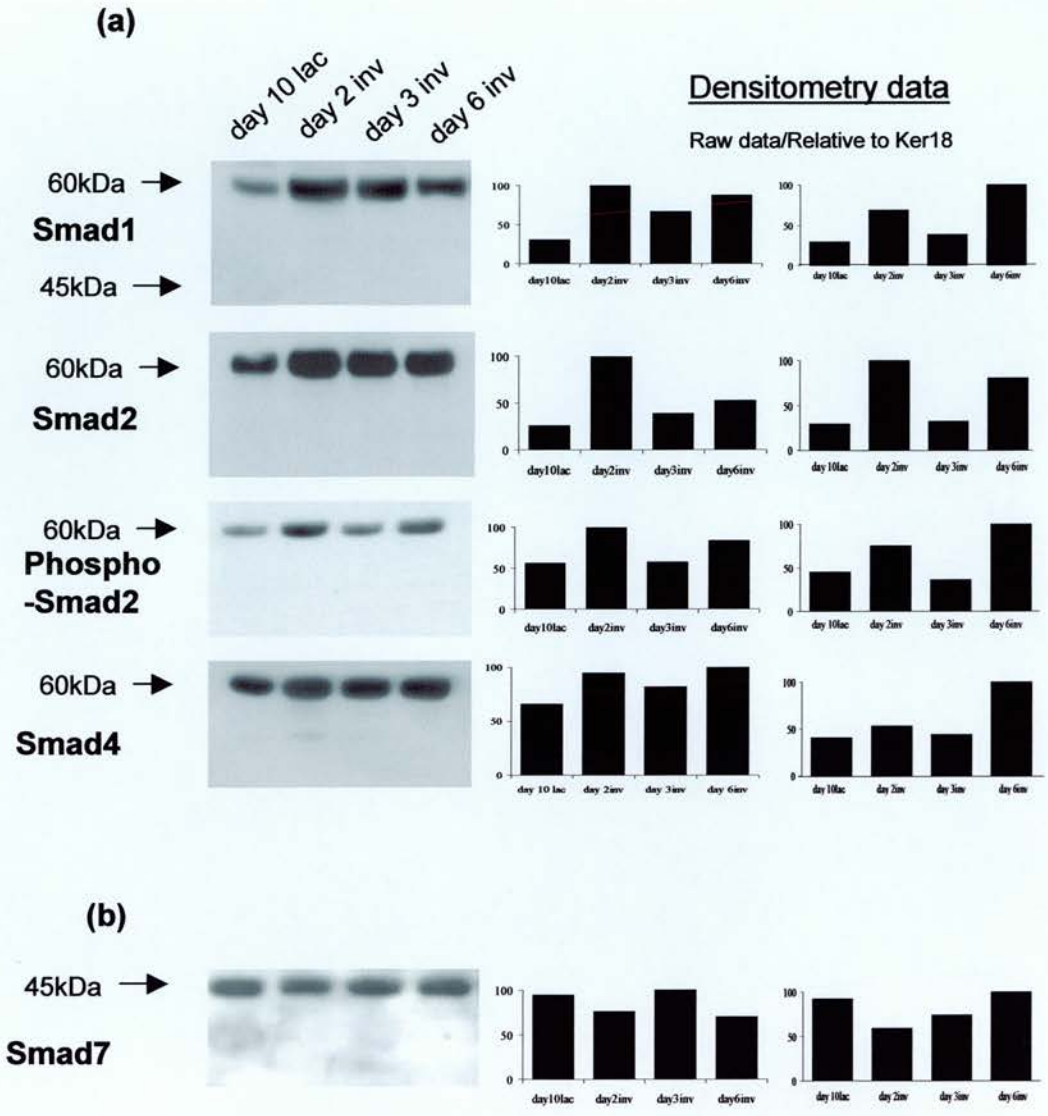


Figure 6.4

Western blot analysis of Smads in wild type mammary gland

(**a**) Western blot analysis of R-Smads1,2, activated Smad2 (phospho-Smad2) and Smad4 in wild type mammary gland from day 10 lactation through to day 6 involution. Densitometry analysis was carried out on an average of two gels and expressed as a percentage of the highest value for each blot

(**b**) Western blot analysis of Anti-Smad, Smad7 in wild-type mammary gland from day 10 lactation through to day 6 involution. Densitometry analysis was carried out on an average of two gels and expressed as a percentage of the highest value for each blot

littermates. This analysis again showed levels of Smad1, 2 and phospho 2 to be increased at day 2, levels of Smad 7 to be reduced at this time point and only relatively minor changes in the levels of Smad4. When normalised to keratin 18 levels and compared to day 10 of lactation, increases in the levels of Smad1, 2, phospho2 and 4 were seen at day 6.

Recently it has been established that Smads and STATs, specifically Smad1 and STAT3, interact *in vitro* in COS-7 cells (Nakashima *et al*, 1999). This interaction takes place in the nucleus and involves the p300 co-activator molecule bound in a ternary complex. I decided to investigate whether this interaction took place *in vivo* during the course of mammary gland involution, and if the interaction could still take place in the absence of STAT3. It was decided to utilise a transgenic strain conditional for STAT3 in the mammary gland to assess the significance of STAT3 in this complex. This model utilises Cre-lox technology to render approximately 80% of the gland null for STAT3 at full lactation (Chapman *et al*, 1999). Immunoprecipitations with p300 were probed for Smad1 and STAT3 in wild type mammary glands and it was found that both proteins can interact with p300 *in vivo* (**Figure 6.5 (I) and (II)**). Immunoprecipitation with Smad1 and subsequent probing with STAT3 implies that the Smad and STAT molecules are bound to the same p300 molecule in a ternary complex, as has already been demonstrated to occur *in vitro* (**Figure 6.5 (III)**). In mice characterised by mammary gland specific deletion of STAT3, immunoprecipitation with p300 and probing with STAT3 confirmed the absence of STAT3 from the p300 complex throughout lactation and involution (**Figure 6.6 (a)**). However, Smad1 was found to bind p300 regardless of the presence or absence of STAT3 (**Figure 6.6 (b)**). To further investigate Smad-STAT interactions I also probed the p300 immunoprecipitations with STAT1 and found that this family member can also bind the p300 molecule irrespective of STAT3 status (**Figure 6.6 (c)**).

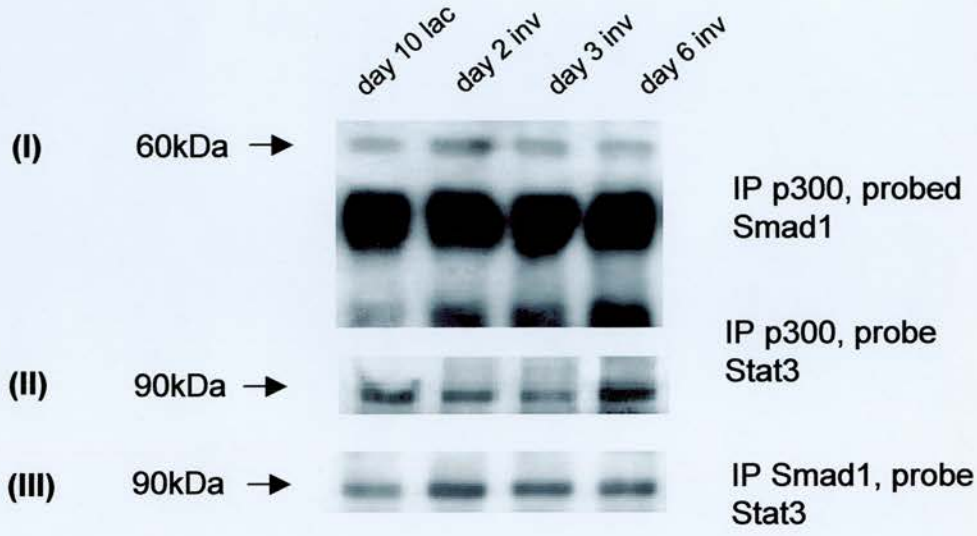


Figure 6.5

Immunoprecipitations with p300 and Smad1, probed with Smad1 and STAT3

(**I**) Immunoprecipitation with p300 on wild type mammary gland tissue at days 10 of
lactation and days 2, 3 and 6 of involution. This blot was probed with Smad1. The large bands at approximately 50kDa and 25kDa reflect binding binding to IgG heavy and light chains and are present in all the IP blots.

(**II**) Immunoprecipitation with p300 on wild type mammary gland tissue at day 10 of lactation and days 2, 3 and 6 of involution, probed with STAT3.

(**III**) Immunoprecipitation with Smad1 on wild type mammary gland tissue at days 10 of lactation and days 2, 3 and 6 of involution, probed with Stat3.

Note: Immunoprecipitation with Smad1, probed with Smad1, confirmed that immunoprecipitations were effective.

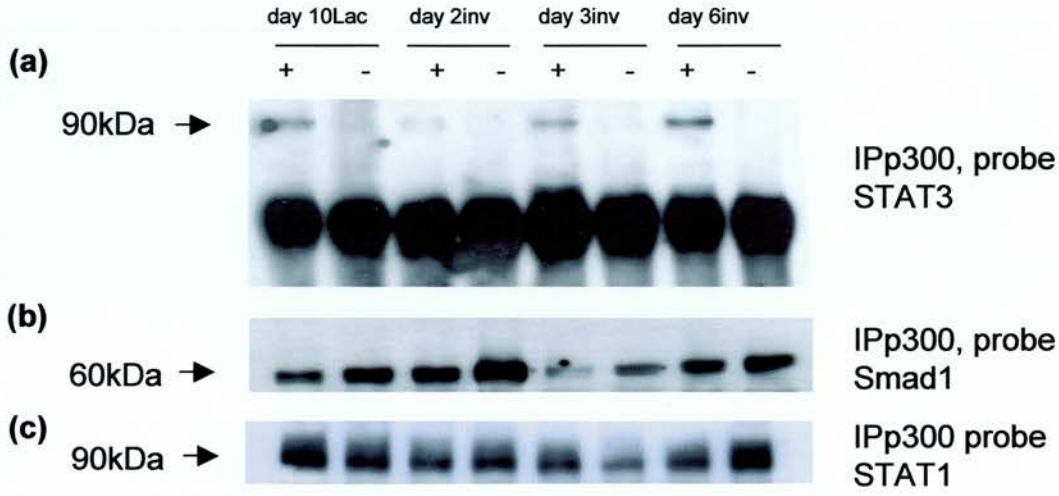


Figure 6.6

Immunoprecipitations with p300 probed with STAT3, Smad1 and STAT1

(**a**) Immunoprecipitation with p300 on wild type (+) and STAT3 null (-) mammary gland tissue at day 10 lactation and days 2, 3 and 6 of involution, probed with a STAT3 antibody. The large bands at approximately 50kDa and 25kDa reflect binding to IgG heavy and light chains and are present in all the IP blots.

(**b**) Immunoprecipitation with p300 probed with Smad1, on wild type (+) and STAT3 null (-) mammary tissues at day 10 lactation and days 2, 3 and 6 of involution.

(**c**) Immunoprecipitation with p300, probed with STAT1 on wild type (+) and STAT3 null (-) mammary tissues at day 10 lactation and days 2, 3 and 6 of involution.

6.2.2 Smad protein levels in a model of delayed involution, a conditional STAT3 knockout

Mammary gland-specific STAT3 null mice have a phenotype of delayed involution (Chapman *et al*, 1999). In these mice the mammary gland involutes approximately three days later than in wild-type littermates and involves up-regulation of genes not associated with normal involution. When levels of the Smads were compared between wild type and STAT3-/- glands, an increase was seen in the absolute levels of Smads1, 2 and phospho2 in the absence of STAT3 at day 3 of involution. No such differences were noted for Smad4 or Smad7 (**Figure 6.7 (a)**). However, these differences must be interpreted in the context of altered involution in the absence of STAT3. As has previously been shown (Chapman *et al*, 1999), STAT3 deficiency causes a delay in involution which leads to a relative increase in the number of epithelial cells during involution. Quantitation of Smad levels relative to the level of keratin 18 (a marker of epithelial cells) suggests that any differences seen between the wild type and the STAT3 null glands at day 2 and day 3 are due to the presence of relatively more epithelial cells in the absence of STAT3 (**Figure 6.7 (b)**). At day 6 of involution, comparison relative to keratin 18 shows a marked increase in the levels of all the Smads in the presence of STAT3.

6.2.3 PAI-1 and p21/WAF1/Cip1 levels in the mammary gland

Several recent studies have used Plasminogen Activator Inhibitor-1 (PAI-1) as a functional readout of Smad1, 2 and 3 activity as the PAI-1 promoter contains multiple Smad binding elements (SBE) which can be activated by these R-Smads (Pearson *et al*, 1999 and Dennler *et al*, 1998). It was therefore considered likely that if Smad-STAT interactions play an important part in the induction and maintenance of involution of the mammary gland then the absence of STAT3 from the p300 complex might alter the functional readout in terms of known downstream gene expression. Consequently I decided to examine PAI-1 protein levels in normal and STAT3-/- mammary glands. PAI-1 levels were compared from wild type and STAT3 -/- mammary glands from day10 of lactation through the first 6 days of involution. It

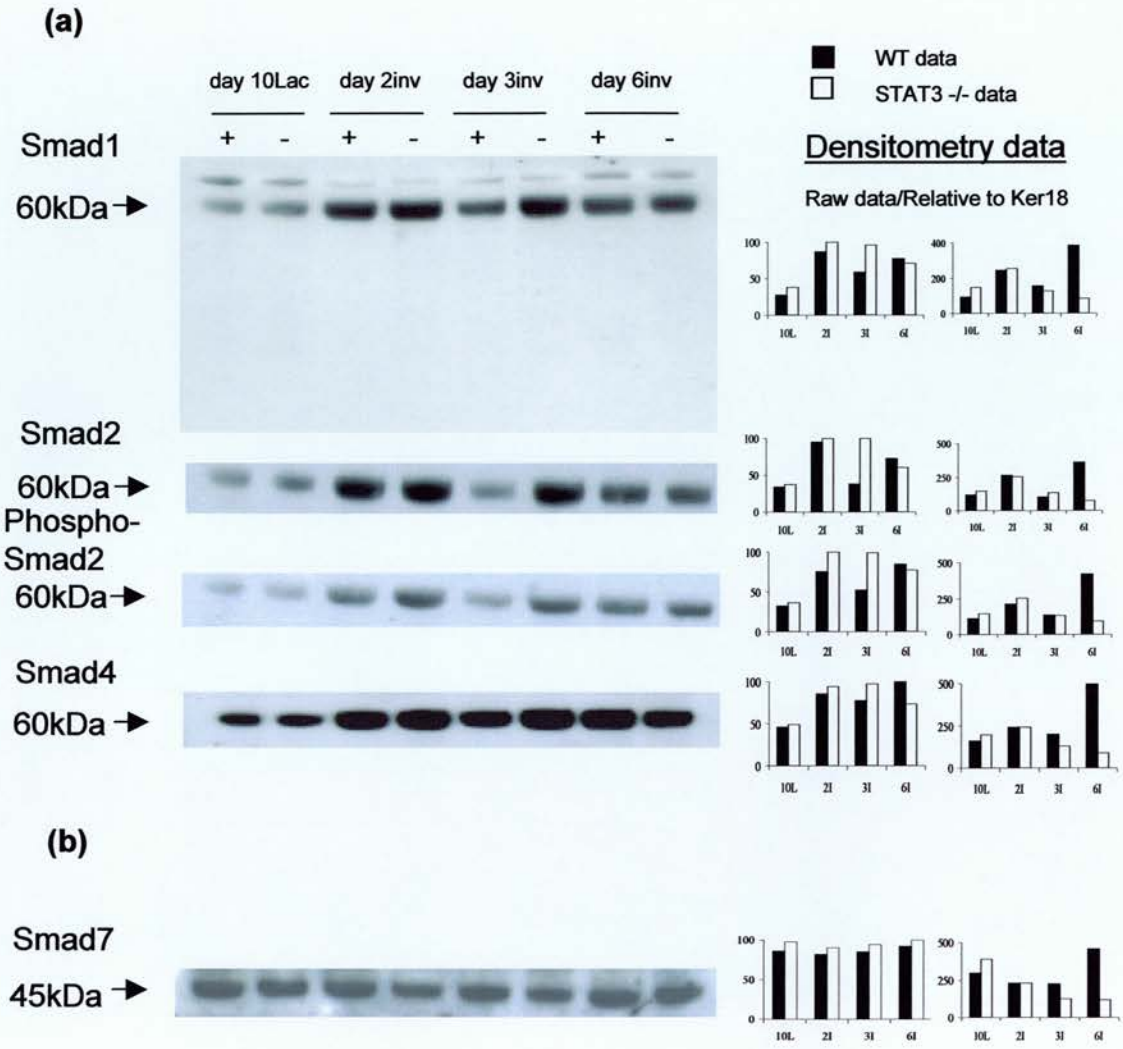


Figure 6.7

(a) Western blot analysis of Smad1, Smad2, phospho-Smad2 and Smad4 in wild type (+) and STAT3 null (-) mammary gland tissues throughout lactation and involution.

Densitometry analysis was carried out on an average of two gels and expressed as a percentage of the highest value for each blot. Solid bars represent wild type and open bars represent STAT3 null. The raw densitometry data was then normalised to keratin 18 data to compensate for epithelial cell content.

(b) Western blot analysis of Smad7 in wild type (+) and STAT3 null (-) mammary gland tissues during lactation and involution.

was found that in normal involuting mammary gland absolute levels of PAI-1 were high at the start of involution, began to decrease by day 3 and were very low at day 6 of involution. In contrast to this the absolute levels of PAI-1 increased throughout involution in the STAT3 null glands to day 3 but dropped by day 6 (**Figure 6.8 (a)**). It is necessary to normalise these results to keratin 18 levels to account for differences in the epithelial cell content of wild type and mutant glands. This normalisation process showed PAI-1 levels to be unaffected by STAT3 deficiency and that even in wild type mammary glands they did not parallel the levels of the Smads, as might be expected if they are accurate downstream markers of gene activation.

To continue with this approach I chose to examine protein levels of p21/WAF1/Cip1, another cited readout of Smad activity. This cell cycle arrest protein is commonly activated in response to TGF-β induced growth arrest and has been shown to be downstream of the Smad4 signalling molecule (Grau *et al*, 1997) with direct evidence for Smad interactions with the p21 promoter region (Moustakas and Kardassis, 1998). Comparison of p21 protein levels in the wild type and STAT3 -/- glands shows that p21 is massively up regulated in the null animals but not in the wild types, both by analysis of raw data and following normalisation to keratin 18 (**Figure 6.8(b)**). Thus, examination of two commonly cited readouts of Smad activity (PAI-1 and p21) has shown that *in vivo* levels of PAI-1 do not parallel levels of the R-Smad proteins as might have been predicted, and that these levels of protein expression are unaffected by STAT3 status. However, these experiments have also shown that p21 levels markedly increase in the absence of STAT3 compared to the very low levels in wild types, and that the peak of p21 expression coincides with the induction of Smads 1, 2 and phospho2 at day 2 of involution.

6.2.4 DNA Binding Activity of Smads in the Mammary Gland

The data obtained from the analysis of PAI-1 and p21/Waf1/Cip1 expression levels therefore gives apparently contradictory readouts of the effects of STAT3 deficiency upon *in vivo* Smad activity in the mammary gland. For both PAI-1 and p21 it is

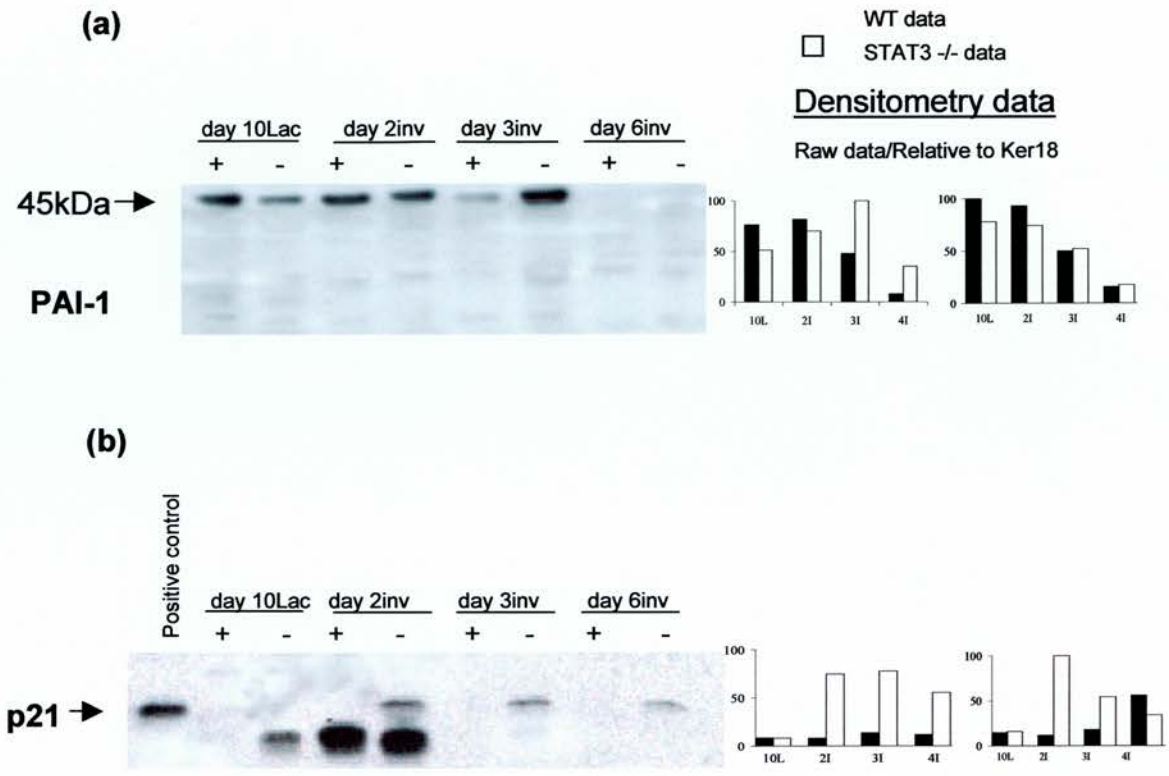


Figure 6.8

(a) Western blot analysis of Plasminogen Activator Inhibitor-1 (PAI-1) in wild type (+) and STAT3 null (-) mammary gland tissues during lactation and involution.

Densitometry analysis was carried out on an average of two gels and expressed as a percentage of the highest value for each blot. Solid bars represent wild type and open bars represent STAT3 null. The raw densitometry data was then normalised to keratin 18 data to compensate for epithelial cell content.

(b) Western blot analysis of p21/Waf-1/Cip1 in wild type (+) and STAT3 null (-) mammary gland tissues during lactation and involution.

Densitometry analysis was carried out on an average of two gels and expressed as a percentage of the highest value for each blot. Solid bars represent wild type and open bars represent STAT3 null. The raw densitometry data was then normalised to keratin 18 data to compensate for epithelial cell content

likely that expression levels during involution of the tissue are affected by factors other than the Smads. Hence, although these are the only currently recognised functional readouts of Smad activity they probably do not accurately reflect Smad activity within the involuting mammary gland. For this reason I decided to examine Smad DNA binding activity and directly determine if this was altered in the absence of STAT3. This question was addressed using an Electrophoretic Mobility Shift Assay (EMSA). The MH1 domains of Smads are known to bind to specific DNA sequences that contain 5'-AGAC-3' sequences, termed Smad Binding Elements (SBEs). EMSA analysis of wild type and STAT3 null mammary gland nuclear extracts showed strong binding to the SBE (GTCTAGAC) in the STAT3 nulls but extremely weak binding in wild type samples. Notably this increased binding was demonstrated for the time point characterised by maximal p21 induction, at day 2 of involution. No binding occurred in either sample with a mutant SBE oligonucleotide demonstrating that this binding was specific to the SBE (**Figure 6.9 (a)**). Supershift analysis with Smad1 and Smad4 showed supershifts with both antibodies in the STAT3 null samples and evidence of a much weaker supershift in the wild type samples (**Figure 6.9 (b)**)

6.3 Discussion

This work has shown that in the normal murine mammary gland Smad1, Smad2 and phospho-Smad2 are all up regulated at the start of involution, both by the analysis of raw data and when calculated relative to keratin 18. Levels of Smad4, the co-Smad, increase only moderately at the start of involution. By contrast, the I-Smad, Smad7, is decreased during this period. These observations suggest that the Smads may be involved in regulating epithelial cell apoptosis which commences at approximately day 2 of involution. Although the changes in protein levels reported here are relatively modest (up to 4-fold induction) these increases are comparable with changes in expression levels known to have important biological consequences. Examples of these include haploinsuffiency of the *Pax* genes (Ostrom *et al,* 2000), the ATM gene (Shigeta *et al*, 1999 and Bay *et al*, 1999) and the *Eya-1* gene, possibly

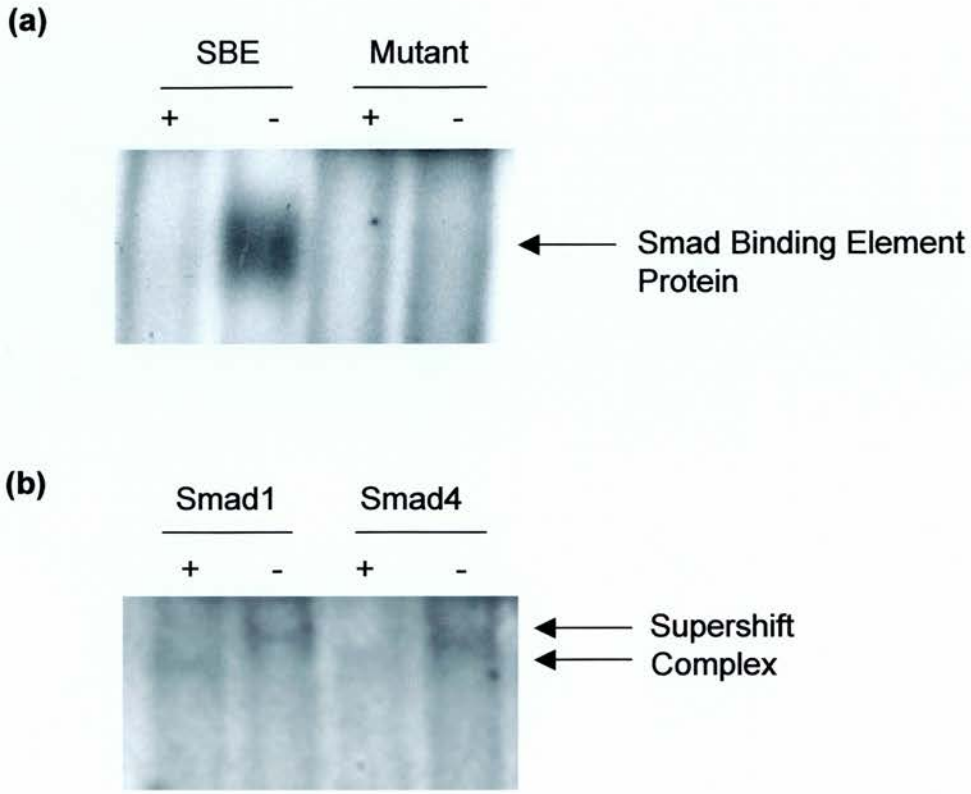


Figure 6.9

Electrophoretic Mobility Shift Analysis (EMSA) of wild type (+) and STAT3 null (-) nuclear extracts from mammary gland tissue at day 2 of involution.

(**a**) Shows part of an EMSA gel where nuclear extracts from wild type (+) and STAT3 null (-) mammary glands have been incubated with either a wild type SBE radiolabelled oligonucleotide or a mutant SBE radiolabelled oligonucleotide.

(**b**) Shows part of an EMSA gel where nuclear extracts from wild type (+) and STAT3 null (-) mammary glands have been incubated with a radiolabelled wild type SBE oligonucleotide plus either Smad2 or Smad4 antibodies.

the most extreme phenotype with haploinsufficient *Eya-1* animals lacking ears and kidneys (Xu *et al*, 1999).

Nakashima and colleagues (Nakashima *et al*, 1999) recently demonstrated Smad1, STAT3 and p300 form a ternary complex *in vitro*. This prompted an investigation to determine if this complex was intact *in vivo* in the mammary gland during lactation and involution. Immunoprecipitation studies show that p300 can bind both STAT3 and Smad1 throughout lactation and involution, strongly suggesting that this complex is intact *in vivo* in wild type mice. Assuming that immunoprecipitations are only semi-quantitative there is no conclusive change in binding to the complex over this period. Immunoprecipitation with Smad1 followed by probing with STAT3 implies that the Smad1 and STAT3 molecules are both bound to the same p300 molecule, indicating that this is probably a tripartite complex with p300 acting as a bridging molecule as has been previously been shown *in vitro* by Nakashima *et al*. The function of this complex remains unknown but Smad1 binding to p300 has recently been shown to synergistically facilitate transcriptional activation (Pearson *et al*, 1999).

The data suggesting that Smads may be involved in the regulation of involution and that Smads and STATs can interact in the mammary gland *in vivo* prompted an enquiry to examine how loss of one of these components, namely STAT3, would affect the process of involution and give insight into the importance of the Smad-STAT complex. The results show that the pattern of Smad protein expression over the course of involution was altered in the absence of STAT3, the principal difference being observed at day 3 of involution. These conclusions are drawn from an analysis of raw densitometric measurement. However, these measurements can be misleading, as the mammary gland is a dynamic tissue which is undergoing massive deconstruction and remodelling during involution. This is particularly relevant to the analysis of STAT3 deficiency as this leads to persistence of the mammary epithelium. It was therefore necessary to normalise all densitometry data against a marker of epithelial cell content, namely Keratin 18, so compensating for variations in epithelial cell content between wild type and STAT3 null glands. This analysis

revealed that, relative to the epithelial cell content, levels of the Smads were unaltered in the absence of STAT3 for the first three days of involution, but that marked differences were observed at day 6. It can be concluded from this that the Smads are not targets of STAT3 in the first three days of involution, and the differences at day 6 must be interpreted with caution because at this time point the vast majority of the gland is composed of non-epithelial cells.

Immunoprecipitation with p300 shows as expected that very little STAT3 binding occurs in the STAT3 null mammary glands. The residual level of binding observed probably reflects the fact that in this model STAT3 deletion occurs only in the epithelial cells of the gland such that some STAT3 will remain in stromal cells. Immunoprecipitation with p300 and subsequent probing with Smad1 shows that Smad1 can, however, bind the p300 molecule in the absence of STAT3. It can be concluded therefore that STAT3 is not absolutely required for the Smad1/p300 complex to form. This also suggests that the p300 molecule is acting as a bridge and that Smad1 and STAT3 may not be directly linked, confirming the data presented by Nakashima *et al* (1999). I have also shown that STAT1 can bind the p300 molecule in both the presence and absence of STAT3. It has been observed that up-regulation of STAT1 protein levels occurs in the absence of STAT3 (see Chapman *et al*, 1999). This raises the possibility that STAT1 may be compensating for STAT3 loss. The immunoprecipitations performed here do not support increased STAT1 binding; however these can only be considered semi-quantitative.

In order to ascertain if functional activity of the Smads was affected by disruption of the p300/Smad/STAT complex I then analysed expression of two commonly cited readouts of Smad expression, namely Plasminogen Activator Inhibitor-1 (PAI-1) and p21/Waf1/Cip1. PAI-1 plays an important part in mammary gland involution as it is one of many protease inhibitors produced by epithelial cells which are down regulated to allow proteolytic degradation of the gland (Zavizion *et al*, 1996). The PAI-1 promoter contains Smad Binding Elements (SBE) in its promoter sequences which have been shown to be targets of transcriptional activation by the Smad proteins. The PAI-1 promoter is commonly used *in vitro* assays to determine Smad2

and 3 activities and very recently was shown to be an appropriate marker for Smad1 activity (Pearson *et al*, 1999). It therefore seemed logical to analyse PAI-1 protein levels as a functional readout of Smad activity in the mammary gland. Once normalised to the keratin 18 content the data shows that in the absence of STAT3 PAI-1 levels remain unaltered relative to wild type levels seen in the glands during involution. Notably, however, changes in PAI-1 expression do not parallel changes in Smad1, 2 and phospho Smad2 expression suggesting that in this context PAI-1 expression is not an accurate indicator of Smad activity. This raises important issues regarding the use of this molecule as a suitable readout *in vivo* for the Smads. It is hardly surprising that in the context of massive mammary gland proteolytic degradation and reconstruction that an inhibitor of proteases is down regulated to allow involution to commence and my findings strongly suggest that there are many more factors other than the Smads acting upon it. A second cited readout of Smad transcriptional activity is p21. Western blot analysis shows that in wild type glands p21 levels remain relatively low and constant throughout lactation and involution. In contrast p21 levels are dramatically increased during involution in the STAT3 null, peaking at day 2 of involution. This pattern of expression only correlates with changes in Smad protein levels in the absence of STAT3 suggesting that STAT3 may normally function by repressing Smad transcriptional activity in the mammary gland.

Use of PAI-1 and p21 protein levels as readouts of Smad activity is potentially flawed because levels of these proteins are likely to be affected by other factors. This may explain the apparently contradictory results obtained for these proteins here. I therefore also investigated if the ability of the Smads to bind target DNA sequences in the mammary gland was altered by disruption of the p300/Smad1/STAT3 complex. EMSA analysis with a consensus SBE oligonucleotide, but not a mutant SBE oligonucleotide, showed that binding occurred much more strongly in the absence of STAT3 than in wild type mammary extracts, supporting the concept that STAT3 can play an inhibitory role in this complex. In order to rule out the possibility that the observed binding to the Smad oligonucleotide was an artefact supershift analysis with anti-Smad antibodies was carried out, confirming that both Smad1 and Smad4 are bound to the complex. The possible mechanism of repression of Smad

binding by STAT3 is currently not known but STAT3 is known to both enhance and repress transcription, one example of which is the effect STAT3 has on the G1 to S transition phase. Here STAT3 has been shown to up regulate cyclins D2, D3 and A whilst concomitantly down regulating p21 and p27. This particular paper suggests that such contradictory signals regulating cell cycle progression could be simultaneously delivered via STAT3 and Smad-STAT interactions could well play a role in orchestrating this kind of balance. (Fukada *et al*, 1998) Another example shows STAT3 repression of c-myc and c-myb and expression of junB and IRF1. In this system blocking of STAT3 activation resulted in inhibition of the repressed c-myb and c-myc (Nakajima *et al*, 1996). Repression of DNA binding can occur via many means including conformational changes and "squelching" – that is sequestration of associated transcriptional proteins. As both Smad and STAT transcriptional complexes are only beginning to be elucidated it remains to be seen which of these possibilities could account for the observed increase in DNA binding by Smads in the absence of STAT3 (**Figure 6.10**).

In conclusion, this work has shown that R-Smads1, 2 and phosphorylated Smad2 are modestly up regulated during involution of the normal mammary gland. The data presented is consistent with the presence of a p300/Smad1/STAT3 complex during mammary gland lactation and involution and further shows that another member of the STAT family, STAT1, is also capable of binding this complex. Analysis of the effect of removing STAT3 from this complex using mice bearing a conditional STAT3 null allele showed that in the absence of STAT3, Smad1 and STAT1 are still capable of binding p300, showing clearly that STAT3 is not required for the binding of these proteins to the complex. STAT3 deficiency does not alter levels of the Smad proteins over the first three days of involution, however protein levels of p21 (one of the transcriptional targets of the Smads) are markedly elevated in the absence of STAT3 during this period, suggesting that STAT3 may be inhibiting Smad activity. By contrast, levels of PAI-1, the only other commonly cited readout of Smad activity are unaffected by STAT3 deficiency and do not parallel the pattern of Smad protein expression. This finding questions the usefulness of PAI-1 as an *in vivo* readout of Smad activity. In order to directly address whether Smad transcriptional activity may

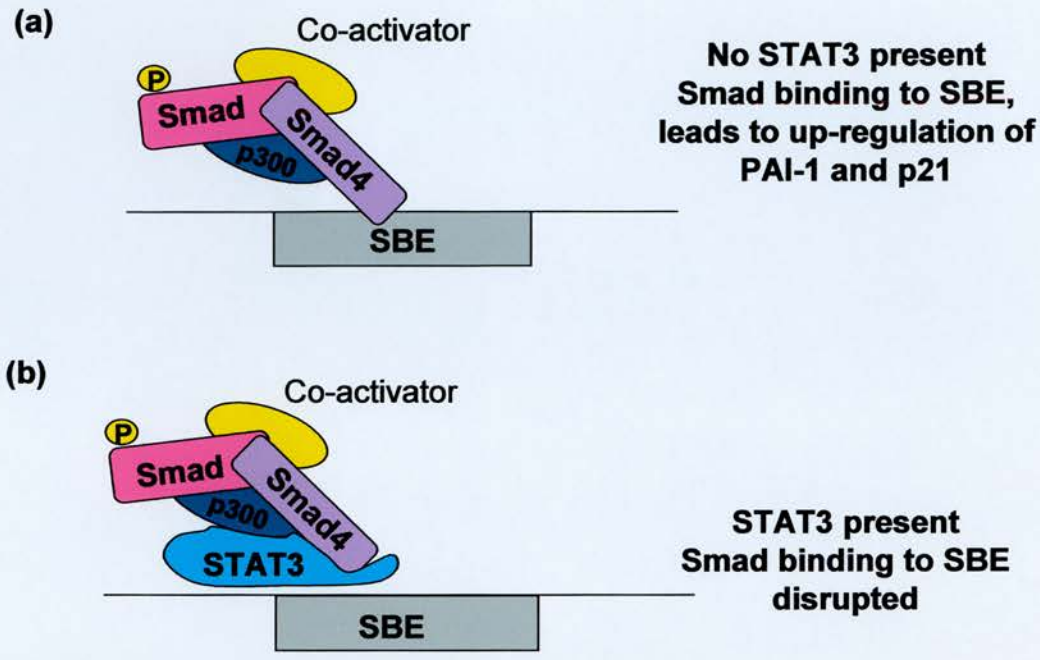


Figure 6.10

Possible model for Smad-STAT Interactions

This figure shows a possible model for Smad-STAT interactions where STAT3 may act as an inhibitor to Smad binding to SBEs within promoter sequences of DNA. This model shows the simplest type of steric inhibition with (**a**) showing that when there is no STAT3 present the complex of R-Smad, Co-Smad, p300 and various co-activators can bind the SBE and initiate transcription of genes such as PAI-1 and p21. In (**b**) the diagram shows how STAT3 may bind preferentially to the p300 molecule thus sterically inhibiting binding of the Smad molecules to the SBE.

be affected by STAT3 status, EMSA analysis was performed upon samples from wild type and STAT3 null glands. This showed that Smad binding to an SBE oligonucleotide was up regulated in the absence of STAT3. Taken together, this data implicate the Smads in regulation of the normal physiology of the mammary gland and show that in this tissue they interact with STAT3, and specifically that STAT3 can mediate repression of Smad DNA binding ability.

Chapter 7 – Discussion

7.1 Introduction

In this thesis I have examined the role of Smad4 in the murine mammary gland by a number of *in vitro* and *in vivo* analyses including the use of two transgenic approaches. I have reported that over-expression of various Smad cDNA expression constructs in murine embryonic stem cells results in both the induction of apoptosis and a G1 arrest; this work was then extended to an examination of the role of the central mediator, Smad4, in the induction of mammary gland epithelial cell apoptosis via the production of both a Smad4 mammary gland transgenic model and a Smad4 conditional deletion. The final analysis was an investigation into Smad-STAT signalling interactions within the mammary gland and an examination on how disruption of one pathway may affect the other.

7.2 Smads and Apoptosis

The initial study to address the question of what role Smad4 might play in mammary gland epithelial cell apoptosis (involution) was based on the known facts that TGF-β plays many important roles in mammary gland physiology including apoptosis during involution, and that Smads had been shown in a number of different cellular environments to be capable of initiating apoptosis. To investigate if different human Smad cDNA sequences were capable of initiating apoptosis in murine ES cells these constructs were transfected and Annexin V levels analysed as an indicator of early apoptotic events. Surprisingly, this experiment showed that only over expression of the central mediator, Smad4 could initiate statistically significant increases in the number of apoptotic cells. This phenomenon had been shown previously where over expression of wild-type Smad4, but not Smad3, in MDCK cells strongly increased the basal level of the p3TP-Lux reporter expression plasmid and initiated apoptosis as determined by TUNEL analysis (Atfi *et al*, 1997). My finding confirms the hypothesis that Smad4 has an important central role to play in the induction of

cellular apoptosis. In an effort to elucidate what mechanism may potentiate this apoptosis I examined cell cycle distributions from ES cells in which the human Smad cDNAs had been over expressed. Previous workers have shown a slight increase in the number of cells in G1 concomitant with Smad4 associated apoptosis (Dai et al, 1999) and I was able to confirm that this also occurred in ES cells. However, this finding must be interpreted with caution as TGF-β induced apoptosis has been associated with many different factors, pathways and causative agents and there is no direct evidence that cell cycle arrest is a causative factor in the induction of apoptosis. Nevertheless, this initial study into the role of Smad over expression prompted an investigation of the *in vivo* role of Smad4 in mediating apoptosis through the production of two transgenic models.

7.3 Transgenic Animal Models

In order to understand what contribution to mammary gland epithelial apoptosis Smad4 might make, two systems were generated to examine over expression and loss of the Smad4 protein in the same cell types of the mammary gland. This approach involved the production of a Smad4 mammary specific transgenic strain and a Smad4 mammary specific conditional deletion strain. Both approaches utilise the same promoter to obtain Smad4 gene expression/deletion with the same cell specificity. I report here the successful creation of both models and, although the questions to be addressed by these remain as yet mostly unanswered, some potential future analyses are discussed.

7.3.1 Creation of the BLG-Smad4 Transgenic

Both models utilise the ovine beta-lactoglobulin (BLG) promoter and it is hoped that the benefits of using this will become apparent when more data is available. The BLG promoter was used to create the Smad4 transgenic and the presence of FLAG-tagged Smad4 protein in the transgenic cohort reflects human Smad4 cDNA over expression. Preliminary analysis suggests these mice show a phenotype of accelerated involution and completion of these analyses together with further

biochemical analysis should be very revealing. The phenotype of the BLG-Smad4 transgenic animals involves no gross abnormalities of the gland which might result in undernourishment of the pups, for example. Instead, initial insights might be drawn from H&Es collected from the transgenic animals, which appear to be further advanced through involution than control mice at the same stage, and it is hoped that further biochemical analyses will confirm this finding. Although these comparisons are made with non-littermate F1 controls, and it remains undetermined whether this observation will hold with F2 littermate control mice, it is unlikely that the use of F1 control animals will alter the final analysis. It is therefore hypothesised that if Smad4 plays an important central role in involution of the gland then over expression of this protein will lead to altered involution. Whatever the final outcome of the phenotype, the model itself will provide valuable information on the precise role of Smad4 during this period.

Determination of expression levels of the human FLAG-tagged Smad4 cDNA will be carried out by Northern blot analysis on the polyA tail of human Smad4 and compared to expression of endogenous mouse Smad4 RNA. There remains a possibility that expression of the transgene is too low to cause a strong phenotype and there are a number of possible reasons as to why this might occur; it has been shown previously that in some cases mammalian cDNAs are not expressed as well as genomic DNAs (Clark *et al*, 1997). For example when the ovine BLG promoter was used to drive expression of some mammalian cDNAs or a CAT reporter sequence it failed to be expressed in the majority of transgenic lines generated. I analysed a total of 10 lines and only one, EK6, expressed the FLAG-tagged transgene. It could well be that this particular cDNA will be poorly expressed in the majority of lines and that a greater number of founder animals would need to be generated in order to overcome this.

Another potential reason for low expression of the transgene might be the absence of a 5'UTR in the BLG promoter (*B. Whitelaw, personal communication*). Although there is currently no evidence to support this idea regarding the efficiency of expression of this particular ovine BLG promoter, it has been shown that 3'UTRs

can aid in the stability of the RNA during processing (Kocarek *et al*, 2000) where the poly(A) tail and sequences contained within the 3' UTR appear to be important for protecting mRNA from RNase activity associated with the translation machinery.

BLG is one of the few promoters which has been shown to be remarkably non-susceptible to position effects (Whitelaw *et al*, 1992). Chromosomal position effects can influence strongly the transcription of foreign genes in transgenic animals and this results in low frequencies and levels of gene expression and, in some cases, in aberrant patterns of expression. It is assumed that transgene arrays integrate into the host genome in a random manner where the integration site often influences the level or pattern of expression of the transgene which may therefore be prone to position effects (Al-Shawi *et al*, 1990). However, expression of heterologous genes from the BLG promoter can result in either no expression (e.g. BLG/CAT as described above) or expression levels unrelated to copy number. In an attempt to overcome this problem, a transgene "rescue" strategy has been devised (Clark *et al*, 1992) based on co-injection with genomic BLG constructs which has been found to result in a higher frequency of expressing lines. This was successful in a number of cases but is not universally applicable suggesting that genomic BLG sequences have the ability to insulate transgenes from position effects in some chromosomal locations but not in others. This "rescue" strategy was not applied to my pronuclear injection sessions (as both were carried out outwith the Roslin Institute) and may be a useful strategy to employ when utilising the BLG promoter in the future.

7.3.2 Future Work with the BLG-Smad4 Transgenic

Future work with this model will be carried out by other workers and will centre on a biochemical analysis of Smad proteins (R-Smads 1, 2 and phospho-2 and the I-Smads 6 and Smad7), downstream targets such as p21 and PAI-1 and STAT proteins, specifically STAT3. Western blot analysis will attempt to determine if over-expression of Smad4 in the mammary gland affects levels of the other Smads, if Smad4 levels are self-limiting, or if R-Smads are the limiting factors. Another possibility is that levels of I-Smads may be up-regulated to help "squelch" activated

R-Smad levels (if indeed these are up-regulated) and in these studies it is hoped to address the question of how central Smad4 is to the propagation and regulation of TGF-β superfamily signalling - the initial experiments presented in this thesis suggest that Smad4 may be the limiting factor in the pathway, and although experiments with Smad4 null cells have shown that the presence of Smad4 is not always necessary for transcription to occur (Dai *et al,* 1999; Sirard *et al*, 2000), and it is possible that Smad4 may not be the only limiting factor in this pathway, this model may help to elucidate how regulation of Smad4 by other Smads might occur.

Analysis of downstream marker genes such as p21and PAI-1 (and associated problems in this analysis as discussed in **Chapter 6**) will provide information on whether over expression of the transgene results directly in increases in transcription. This study will enable us to address the question of whether other nuclear signals might be involved (for example nuclear transcriptional co-repressors like TGIF (Wotton *et al*, 1999) although it is currently difficult to assess exactly which molecules due to lack of availability of antibodies.

Analysis of STAT protein levels follows on from my observations (in **Chapter 6**) that Smads and STATs can interact in the context of mammary gland lactation and involution and that STAT3 may be inhibitory to the Smad1 transcriptional apparatus. If this phenomenon of STAT inhibition is truly a mechanism of Smad transcriptional control then it might be expected that a STAT protein could be up regulated to compensate for the increase in Smad signals resulting from the transgene.

It is hoped that the transgenic cohort can be rapidly expanded to the required number of animals such that these experiments and others can help elucidate what role Smad4 plays in mammary gland lactation and involution. However, this question can be addressed in parallel using the converse animal model, that is a conditional deletion of Smad4 in the mammary gland.

7.3.3 Creating the Conditional allele of Smad4

I have reported in this thesis the successful germ line transmission of a conditional allele of the murine Smad4 gene. This method of analysis of Smad4 function in the mammary gland was chosen based on the knowledge that conventional Smad4 gene targeting has resulted in an embryonic lethal phenotype (Sirard *et al*, 1998). Therefore, in order to study loss of Smad4 in the gland this approach was deemed highly appropriate. One factor, which strengthened the argument to embark on such a potentially risky and lengthy project, was the immediate availability of well-characterised mammary gland Cre transgenic animals for mating to the "floxed" mouse (Selbert *et al*, 1998).

7.3.4 Future work with the Conditional "Floxed" Smad4 Model

Due to time constraints the analysis of loss of Smad4 in murine mammary gland epithelium lies outwith this project. The immediate next step is to mate the floxed Smad4 mice (with one wild type allele and one "floxed" allele, termed $Smad4^{+/fl}$) to animals heterozygous for Smad4 ($Smad4^{+/-}$) (Sirard *et al*, 1998) to create animals with one floxed allele and one null allele ($Smad4^{fl/-}$). The second step is to mate the $Smad4^{fl/-}$ mice to our BLG-Cre transgenic animals (Selbert *et al*, 1998). These animals have already been demonstrated to express Cre from late gestation until late involution and, as soon as progeny are available, it is hoped to begin characterisation. This will comprise firstly, the kinetics and degree of deletion of Smad4 occurring in the gland and secondly, the phenotype of loss of Smad4 in the murine mammary gland. Analysis of the phenotype will concentrate on histological and wholemount analysis of the mammary glands and on a biochemical analysis of related and downstream target genes, similar to those used to analyse the phenotype in the Smad4 transgenic animal model.

7.4 Smad-STAT Interactions in the Mammary Gland

I have shown here for the first time that the STAT family of signal transduction factors/transcription factors can inhibit Smad initiation of transcription in the mammary gland, and confirm that this inhibition occurs through a complex formed with the co-activator/repressor molecule p300. I have also shown that the commonly cited readout of Smad activity, PAI-1, is not entirely appropriate in analysis of the mammary gland *in vivo* as PAI-1 is likely to be a target of many of the protease pathways active during the involution stage of mammary gland progression. This fact also brings into question the suitability of using p21 as a marker of Smad activity *in vivo* as this molecule is undoubtedly also a target of many of the active signalling pathways at this time point. Studies in progress in collaborative laboratories where mammary-specific floxed STAT3 mice have been crossed to p53 null animals will determine if p21 is indeed a suitable marker for Smads *in vivo*. The fact that these two molecules are the only commonly cited targets of Smad activation has made this analysis difficult but the combination of this and EMSA analysis with an SBE oligonucleotide strongly suggests that STAT3 is indeed inhibitory to Smads in the context of mammary gland involution, and that this inhibition is generated through a complex containing Smad1, STAT3, p300 and probably Smad4.

7.5 Conclusions

In this thesis I have attempted to describe an analysis of the role of Smad4 in the murine mammary gland. This analysis centred on the use of transgenic technology to create potentially useful models of both over expression and loss of Smad4 in the same types of cell within the gland, the epithelial cells. Other analyses have included an *in vitro* investigation into how Smad4 can cause cell death when inappropriately expressed in embryonic stem cells, and an *in vivo* analysis of Smad-STAT interactions within the mammary gland, which has shown inhibition of Smads by STATs for the first time. Characterisation of the animal models will be a lengthy process which lies outwith the realms of a 3-year Ph.D. However, in the near future I

hope that information can be gained from both models which will be constructive in determining what role Smad4 plays in mammary gland lactation and involution.

The evidence presented in this thesis allows several conclusions to be drawn on the role of Smad4 in the induction of apoptosis. These conclusions are; first, that Smad4 is capable of inducing cellular apoptosis (as determined by transient transfection experiments in embryonic stem cells) and that this induction of apoptosis is associated with a G1 arrest; second, that Smad4 levels are up-regulated during the normal programme of apoptosis associated with mammary gland involution and that this up-regulation is concomitant with an increase in the expression of the cell cycle cdk (cyclin dependent kinase) inhibitor p21/Waf1/Cip1, known to inhibit all cdks involved in theG1 $\rightarrow$ S transition. This data strongly supports the notion that p21 is a direct target of the Smad pathway and that altered p21 expression can induce cell death. Finally, preliminary analysis of transgenic mice expressing Smad4 specifically within the mammary epithelium has suggested that this may result in accelerated involution compared to wild type controls. Taken together, this data strongly suggests that Smad4 can modulate the programme of apoptosis and that it plays an important role in both the initiation and maintenance of mammary gland involution.

Bibliography

Abremski, K. and R. Hoess. 1984. Bacteriophage P1 site-specific recombination. Purification and properties of the Cre recombinase protein. J Biol Chem **259**:1509-1514.

Adams, J.M. and S. Cory. 1991. Transgenic models of tumor development. Science **254**:1161-1167.

Afrakhte, M., A. Moren, S. Jossan, S. Itoh, K. Sampath, B. Westermark, C.H. Heldin, N.E. Heldin, and P. ten Dijke. 1998. Induction of inhibitory Smad6 and Smad7 mRNA by TGF-beta family members. Biochem Biophys Res Commun **249**:505-511.

Akira, S. 1999. Functional roles of STAT family proteins: lessons from knockout mice. Stem Cells **17**:138-146.

Akiyoshi, S., H. Inoue, J. Hanai, K. Kusanagi, N. Nemoto, K. Miyazono, and M. Kawabata. 1999. c-Ski acts as a transcriptional co-repressor in transforming growth factor-beta signaling through interaction with smads. J Biol Chem **274**:35269-35277.

al-Shawi, R., J. Kinnaird, J. Burke, and J.O. Bishop. 1990. Expression of a foreign gene in a line of transgenic mice is modulated by a chromosomal position effect. Mol Cell Biol **10**:1192-1198.

Alleva, D.G., C.J. Burger, and K.D. Elgert. 1994. Tumor-induced regulation of suppressor macrophage nitric oxide and TNF-alpha production. Role of tumor-derived IL-10, TGF-beta, and prostaglandin E2. J Immunol **153**:1674-1686.

Anderson, W.F., L. Killos, L. Sanders-Haigh, P.J. Kretschmer, and E.G. Diacumakos. 1980. Replication and expression of thymidine kinase and human globin genes microinjected into mouse fibroblasts. Proc Natl Acad Sci U S A **77**:5399-5403.

Anton, M. and F.L. Graham. 1995. Site-specific recombination mediated by an adenovirus vector expressing the Cre recombinase protein: a molecular switch for control of gene expression. J Virol **69**:4600-4606.

Archibald, A.L., M. McClenaghan, V. Hornsey, J.P. Simons, and A.J. Clark. 1990. High-level expression of biologically active human alpha 1-antitrypsin in the milk of transgenic mice. Proc Natl Acad Sci U S A **87**:5178-5182.

Ariazi, E.A., Y. Satomi, M.J. Ellis, J.D. Haag, W. Shi, C.A. Sattler, and M.N. Gould. 1999. Activation of the transforming growth factor beta signaling pathway and induction of cytostasis and apoptosis in mammary carcinomas treated with the anticancer agent perillyl alcohol. Cancer Res **59**:1917-1928.

Arora, K., H. Dai, S.G. Kazuko, J. Jamal, M.B. O'Connor, A. Letsou, and R. Warrior. 1995. The Drosophila schnurri gene acts in the Dpp/TGF beta signaling pathway and encodes a transcription factor homologous to the human MBP family. Cell **81**:781-790.

Ashley, D.M., F.M. Kong, D.D. Bigner, and L.P. Hale. 1998. Endogenous expression of transforming growth factor beta1 inhibits growth and tumorigenicity and enhances Fas-mediated apoptosis in a murine high-grade glioma model. Cancer Res **58**:302-309.

Atfi, A., M. Buisine, A. Mazars, and C. Gespach. 1997. Induction of apoptosis by DPC4, a transcriptional factor regulated by transforming growth factor-beta through stress-activated

protein kinase/c-Jun N-terminal kinase (SAPK/JNK) signaling pathway. J Biol Chem **272**:24731-24734.

Attisano, L. and J.L. Wrana. 1996. Signal transduction by members of the transforming growth factor-beta superfamily. Cytokine Growth Factor Rev **7**:327-339.

Attisano, L., J.L. Wrana, E. Montalvo, and J. Massague. 1996. Activation of signalling by the activin receptor complex. Mol Cell Biol **16**:1066-1073.

Austin, S., M. Ziese, and N. Sternberg. 1981. A novel role for site-specific recombination in maintenance of bacterial replicons. Cell **25**:729-736.

Azuma, M., K. Motegi, K. Aota, T. Yamashita, H. Yoshida, and M. Sato. 1999. TGF-beta1 inhibits NF-kappaB activity through induction of IkappaB-alpha expression in human salivary gland cells: a possible mechanism of growth suppression by TGF-beta1. Exp Cell Res **250**:213-222.

Baker, J.C. and R.M. Harland. 1996. A novel mesoderm inducer, Madr2, functions in the activin signal transduction pathway. Genes Dev **10**:1880-1889.

Baker, J.C. and R.M. Harland. 1997. From receptor to nucleus: the Smad pathway. Curr Opin Genet Dev **7**:467-473.

Bannister, A.J. and T. Kouzarides . 1996. The CBP co-activator is a histone acetyltransferase. Nature **384**:641-643.

Basler, K. and G. Struhl. 1994. Compartment boundaries and the control of Drosophila limb pattern by hedgehog protein [see comments]. Nature **368**:208-214.

Bay, J.O., N. Uhrhammer, D. Pernin, N. Presneau, A. Tchirkov, M. Vuillaume, V. Laplace, M. Grancho, P. Verrelle, J. Hall, and Y.J. Bignon. 1999. High incidence of cancer in a family segregating a mutation of the ATM gene: possible role of ATM heterozygosity in cancer. Hum Mutat **14**:485-492.

Behringer, R.R. 1994. The in vivo roles of mullerian-inhibiting substance. Curr Top Dev Biol **29**:171-187.

Behringer, R.R., M.J. Finegold, and R.L. Cate. 1994. Mullerian-inhibiting substance function during mammalian sexual development. Cell **79**:415-425.

Behringer, R.R. 1994. The in vivo roles of mullerian-inhibiting substance. Curr Top Dev Biol **29**:171-187.

Bellone, G., M. Aste-Amezaga, G. Trinchieri, and U. Rodeck. 1995. Regulation of NK cell functions by TGF-beta 1. J Immunol **155**:1066-1073.

Berg, J.N., C.J. Gallione, T.T. Stenzel, D.W. Johnson, W.P. Allen, C.E. Schwartz, C.E. Jackson, M.E. Porteous, and D.A. Marchuk. 1997. The activin receptor-like kinase 1 gene: genomic structure and mutations in hereditary hemorrhagic telangiectasia type 2. Am J Hum Genet **61**:60-67.

Berns, A. 1991. Tumorigenesis in transgenic mice: identification and characterization of synergizing oncogenes. J Cell Biochem **47**:130-135.

Bonyadi, M., S.A. Rusholme, F.M. Cousins, H.C. Su, C.A. Biron, M. Farrall, and R.J. Akhurst. 1997. Mapping of a major genetic modifier of embryonic lethality in TGF beta 1 knockout mice. Nat Genet **15**:207-211.

Bostrom, M.P. and P. Asnis. 1998. Transforming growth factor beta in fracture repair. Clin Orthop S124-S131

Bouchard, C., W.H. Fridman, and C. Sautes. 1997. Effect of TGF-beta1 on cell cycle regulatory proteins in LPS-stimulated normal mouse B lymphocytes. J Immunol **159**:4155-4164.

Boyd, F.T. and J. Massague. 1989. Transforming growth factor-beta inhibition of epithelial cell proliferation linked to the expression of a 53-kDa membrane receptor. J Biol Chem **264**:2272-2278.

Bradley, A., M. Evans, M.H. Kaufman, and E. Robertson. 1984. Formation of germ-line chimaeras from embryo-derived teratocarcinoma cell lines. Nature **309**:255-256.

Brand, T. and M.D. Schneider. 1996. Transforming growth factor-beta signal transduction. Circ Res **78**:173-179.

Brandon, E.P., R.L. Idzerda, and G.S. McKnight. 1995. Targeting the mouse genome: a compendium of knockouts (Part III) [published erratum appears in Curr Biol 1995 Sep 1;5(9):1073]. Curr Biol **5**:873-881.

Brandon, E.P., R.L. Idzerda, and G.S. McKnight. 1995. Targeting the mouse genome: a compendium of knockouts (Part II) [published erratum appears in Curr Biol 1995 Sep 1;5(9):1073]. Curr Biol **5**:758-765.

Brandon, E.P., R.L. Idzerda, and G.S. McKnight. 1995. Knockouts. Targeting the mouse genome: a compendium of knockouts (Part I) [published erratum appears in Curr Biol 1995 Sep 1;5(9):1073]. Curr Biol **5**:625-634.

Brinster, R.L., H.Y. Chen, M. Trumbauer, A.W. Senear, R. Warren, and R.D. Palmiter. 1981. Somatic expression of herpes thymidine kinase in mice following injection of a fusion gene into eggs. Cell **27**:223-231.

Brinster, R.L. and R.D. Palmiter. 1984. Transgenic mice containing growth hormone fusion genes. Philos Trans R Soc Lond B Biol Sci **307**:309-312.

Brinster, R.L. and R.D. Palmiter. 1984. Introduction of genes into the germ line of animals. Harvey Lect **80**:1-38.

Brinster, R.L., H.Y. Chen, M.E. Trumbauer, M.K. Yagle, and R.D. Palmiter. 1985. Factors affecting the efficiency of introducing foreign DNA into mice by microinjecting eggs. Proc Natl Acad Sci U S A **82**:4438-4442.

Brinster, R.L., J.M. Allen, R.R. Behringer, R.E. Gelinas, and R.D. Palmiter. 1988. Introns increase transcriptional efficiency in transgenic mice. Proc Natl Acad Sci U S A **85**:836-840.

Brinster, R.L., E.P. Sandgren, R.R. Behringer, and R.D. Palmiter. 1989. No simple solution for making transgenic mice [letter; comment]. Cell **59**:239-241.

Brinster, R.L., R.E. Braun, D. Lo, M.R. Avarbock, F. Oram, and R.D. Palmiter. 1989. Targeted correction of a major histocompatibility class II E alpha gene by DNA microinjected into mouse eggs. Proc Natl Acad Sci U S A **86**:7087-7091.

Brisken, C., A. Heineman, T. Chavarria, B. Elenbaas, J. Tan, S.K. Dey, J.A. McMahon, A.P. McMahon, and R.A. Weinberg. 2000. Essential function of Wnt-4 in mammary gland development downstream of progesterone signaling. Genes Dev **14**:650-654.

Brodin, G., P. ten Dijke, K. Funa, C.H. Heldin, and M. Landstrom. 1999. Increased smad expression and activation are associated with apoptosis in normal and malignant prostate after castration. Cancer Res **59**:2731-2738.

Brummel, T.J., V. Twombly, G. Marques, J.L. Wrana, S.J. Newfeld, L. Attisano, J. Massague, M.B. O'Connor, and W.M. Gelbart. 1994. Characterization and relationship of Dpp receptors encoded by the saxophone and thick veins genes in Drosophila. Cell **78**:251-261.

Bursch, W., F. Oberhammer, R.L. Jirtle, M. Askari, R. Sedivy, B. Grasl-Kraupp, A.F. Purchio, and R. Schulte-Hermann. 1993. Transforming growth factor-beta 1 as a signal for induction of cell death by apoptosis. Br J Cancer **67**:531-536.

Campbell, G., T. Weaver, and A. Tomlinson. 1993. Axis specification in the developing Drosophila appendage: the role of wingless, decapentaplegic, and the homeobox gene aristaless. Cell **74**:1113-1123.

Capecchi, M.R. 1980. High efficiency transformation by direct microinjection of DNA into cultured mammalian cells. Cell **22**:479-488.

Capovilla, M., M. Brandt, and J. Botas. 1994. Direct regulation of decapentaplegic by Ultrabithorax and its role in Drosophila midgut morphogenesis. Cell **76**:461-475.

Carcamo, J., A. Zentella, and J. Massague. 1995. Disruption of transforming growth factor beta signaling by a mutation that prevents transphosphorylation within the receptor complex. Mol Cell Biol **15**:1573-1581.

Catlett-Falcone, R., T.H. Landowski, M.M. Oshiro, J. Turkson, A. Levitzki, R. Savino, G. Ciliberto, L. Moscinski, J.L. Fernandez-Luna, G. Nunez, W.S. Dalton, and R. Jove. 1999. Constitutive activation of Stat3 signaling confers resistance to apoptosis in human U266 myeloma cells. Immunity **10**:105-115.

Chang, H., D. Huylebroeck, K. Verschueren, Q. Guo, M.M. Matzuk, and A. Zwijsen. 1999. Smad5 knockout mice die at mid-gestation due to multiple embryonic and extraembryonic defects. Development **126**:1631-1642.

Chapman, R.S., P.C. Lourenco, E. Tonner, D.J. Flint, S. Selbert, K. Takeda, S. Akira, A.R. Clarke, and C.J. Watson. 1999. Suppression of epithelial apoptosis and delayed mammary gland involution in mice with a conditional knockout of Stat3. Genes Dev **13**:2604-2616.

Chen, R.H., P.J. Miettinen, E.M. Maruoka, L. Choy, and R. Derynck. 1995. A WD-domain protein that is associated with and phosphorylated by the type II TGF-beta receptor. Nature **377**:548-552.

Chen, W., T.K. Woodruff, and K.E. Mayo. 2000. Activin A-induced HepG2 liver cell apoptosis: involvement of activin receptors and smad proteins. Endocrinology **141**:1263-1272.

Chen, X., M.J. Rubock, and M. Whitman. 1996. A transcriptional partner for MAD proteins in TGF-beta signalling [published erratum appears in Nature 1996 Dec 19-26;384(6610):648]. Nature **383**:691-696.

Chen, X., E. Weisberg, V. Fridmacher, M. Watanabe, G. Naco, and M. Whitman. 1997. Smad4 and FAST-1 in the assembly of activin-responsive factor. Nature **389**:85-89.

Chen, Y.G., A. Hata, R.S. Lo, D. Wotton, Y. Shi, N. Pavletich, and J. Massague. 1998. Determinants of specificity in TGF-beta signal transduction. Genes Dev **12**:2144-2152.

Chiao, P.J., K.K. Hunt, A.M. Grau, A. Abramian, J. Fleming, W. Zhang, T. Breslin, J.L. Abbruzzese, and D.B. Evans . 1999. Tumor suppressor gene Smad4/DPC4, its downstream target genes, and regulation of cell cycle. Ann N Y Acad Sci **880**:31-37.

Christian, J.L. and T. Nakayama. 1999. Can't get no SMADisfaction: Smad proteins as positive and negative regulators of TGF-beta family signals. Bioessays **21**:382-390.

Chrivia, J.C., R.P. Kwok, N. Lamb, M. Hagiwara, M.R. Montminy, and R.H. Goodman. 1993. Phosphorylated CREB binds specifically to the nuclear protein CBP. Nature **365**:855-859.

Clark, A.J., A. Cowper, R. Wallace, G. Wright, and J.P. Simons. 1992. Rescuing transgene expression by co-integration. Biotechnology (N Y) **10**:1450-1454.

Clark, A.J., G. Harold, and F.E. Yull. 1997. Mammalian cDNA and prokaryotic reporter sequences silence adjacent transgenes in transgenic mice. Nucleic Acids Res **25**:1009-1014.

Clark, A.J. 1998. Gene expression in the mammary glands of transgenic animals. Biochem Soc Symp **63**:133-140.

Coerver, K.A., T.K. Woodruff, M.J. Finegold, J. Mather, A. Bradley, and M.M. Matzuk. 1996. Activin signaling through activin receptor type II causes the cachexia-like symptoms in inhibin-deficient mice. Mol Endocrinol **10**:534-543.

Collignon, J., I. Varlet, and E.J. Robertson. 1996. Relationship between asymmetric nodal expression and the direction of embryonic turning [see comments]. Nature **381**:155-158.

Conlon, F.L., K.S. Barth, and E.J. Robertson. 1991. A novel retrovirally induced embryonic lethal mutation in the mouse: assessment of the developmental fate of embryonic stem cells homozygous for the 413.d proviral integration. Development **111**:969-981.

Costantini, F. and E. Lacy. 1981. Introduction of a rabbit beta-globin gene into the mouse germ line. Nature **294**:92-94.

Cunha, G.R. 1994. Role of mesenchymal-epithelial interactions in normal and abnormal development of the mammary gland and prostate. Cancer **74**:1030-1044.

Dai, J.L., R.K. Bansal, and S.E. Kern. 1999. G1 cell cycle arrest and apoptosis induction by nuclear Smad4/Dpc4: phenotypes reversed by a tumorigenic mutation. Proc Natl Acad Sci U S A **96**:1427-1432.

Dai, J.L., M. Schutte, R.K. Bansal, R.E. Wilentz, A.Y. Sugar, and S.E. Kern. 1999. Transforming growth factor-beta responsiveness in DPC4/SMAD4-null cancer cells. Mol Carcinog **26**:37-43.

Dai, J.L., R.K. Bansal, and S.E. Kern. 1999. G1 cell cycle arrest and apoptosis induction by nuclear Smad4/Dpc4: phenotypes reversed by a tumorigenic mutation. Proc Natl Acad Sci U S A **96**:1427-1432.

Daniel, C.W., G.B. Silberstein, and P. Strickland. 1987. Direct action of 17 beta-estradiol on mouse mammary ducts analyzed by sustained release implants and steroid autoradiography. Cancer Res **47**:6052-6057.

Daniel, C.W. and S.D. Robinson. 1992. Regulation of mammary growth and function by TGF-beta. Mol Reprod Dev **32**:145-151.

de Caestecker, M.P., P. Hemmati, S. Larisch-Bloch, R. Ajmera, A.B. Roberts, and R.J. Lechleider. 1997. Characterization of functional domains within Smad4/DPC4. J Biol Chem **272**:13690-13696.

de Caestecker, M.P., W.T. Parks, C.J. Frank, P. Castagnino, D.P. Bottaro, A.B. Roberts, and R.J. Lechleider. 1998. Smad2 transduces common signals from receptor serine-threonine and tyrosine kinases. Genes Dev **12**:1587-1592.

de Caestecker, M.P., T. Yahata, D. Wang, W.T. Parks, S. Huang, C.S. Hill, T. Shioda, A.B. Roberts, and R.J. Lechleider. 2000. The Smad4 activation domain (SAD) is a proline-rich, p300-dependent transcriptional activation domain. J Biol Chem **275**:2115-2122.

de Winter, J.P., C.J. de Vries, T.A. van Achterberg, R.F. Ameerun, A. Feijen, H. Sugino, P. de Waele, D. Huylebroeck, K. Verschueren, and A.J. Van Den Eijden-Van Raaij. 1996. Truncated activin type II receptors inhibit bioactivity by the formation of heteromeric complexes with activin type I. receptors. Exp Cell Res **224**:323-334.

DeClerck, Y.A. and W.E. Laug. 1996. Cooperation between matrix metalloproteinases and the plasminogen activator-plasmin system in tumor progression. Enzyme Protein **49**:72-84.

Demeneix, B.A., H. Abdel-Taweb, C. Benoist, I. Seugnet, and J.P. Behr. 1994. Temporal and spatial expression of lipospermine-compacted genes transferred into chick embryos in vivo. Biotechniques **16**:496-501.

Deng, C. and M.R. Capecchi. 1992. Reexamination of gene targeting frequency as a function of the extent of homology between the targeting vector and the target locus. Mol Cell Biol **12**:3365-3371.

Dennler, S., S. Itoh, D. Vivien, P. ten Dijke, S. Huet, and J.M. Gauthier. 1998. Direct binding of Smad3 and Smad4 to critical TGF beta-inducible elements in the promoter of human plasminogen activator inhibitor-type 1 gene. EMBO J **17**:3091-3100.

Derynck, R., J.A. Jarrett, E.Y. Chen, D.H. Eaton, J.R. Bell, R.K. Assoian, A.B. Roberts, M.B. Sporn, and D.V. Goeddel. 1985. Human transforming growth factor-beta complementary DNA sequence and expression in normal and transformed cells. Nature **316**:701-705.

Derynck, R., W.M. Gelbart, R.M. Harland, C.H. Heldin, S.E. Kern, J. Massague, D.A. Melton, M. Mlodzik, R.W. Padgett, A.B. Roberts, J. Smith, G.H. Thomsen, B. Vogelstein, and X.F. Wang. 1996. Nomenclature: vertebrate mediators of TGFbeta family signals [letter]. Cell **87**:173-173.

Derynck, R. and Y. Zhang. 1996. Intracellular signalling: the mad way to do it. Curr Biol **6**:1226-1229.

Derynck, R. 1998. SMAD proteins and mammalian anatomy [news; comment]. Nature **393**:737-739.

Derynck, R., Y. Zhang, and X.H. Feng. 1998. Smads: transcriptional activators of TGF-beta responses. Cell **95**:737-740.

Dietrich, W.F., J. Miller, R. Steen, M.A. Merchant, D. Damron-Boles, Z. Husain, R. Dredge, M.J. Daly, K.A. Ingalls, and T.J. O'Connor. 1996. A comprehensive genetic map of the mouse genome [see comments] [published erratum appears in Nature 1996 May 9;381(6578):172]. Nature **380**:149-152.

Doetschman, T., R.G. Gregg, N. Maeda, M.L. Hooper, D.W. Melton, S. Thompson, and O. Smithies. 1987. Targetted correction of a mutant HPRT gene in mouse embryonic stem cells. Nature **330**:576-578.

Dou, Q.P., B. An, and P.L. Will. 1995. Induction of a retinoblastoma phosphatase activity by anticancer drugs accompanies p53-independent G1 arrest and apoptosis. Proc Natl Acad Sci U S A **92**:9019-9023.

Duff, E.K. and A.R. Clarke. 1998. Smad4 (DPC4)--a potent tumour suppressor? Br J Cancer **78**:1615-1619.

Eckner, R., Z. Arany, M. Ewen, W. Sellers, and D.M. Livingston. 1994. The adenovirus E1A-associated 300-kD protein exhibits properties of a transcriptional coactivator and belongs to an evolutionarily conserved family. Cold Spring Harb Symp Quant Biol **59**:85-95.

Eppert, K., S.W. Scherer, H. Ozcelik, R. Pirone, P. Hoodless, H. Kim, L.C. Tsui, B. Bapat, S. Gallinger, I.L. Andrulis, G.H. Thomsen, J.L. Wrana, and L. Attisano. 1996. MADR2 maps to 18q21 and encodes a TGFbeta-regulated MAD-related protein that is functionally mutated in colorectal carcinoma. Cell **86**:543-552.

Estevez, M., L. Attisano, J.L. Wrana, P.S. Albert, J. Massague, and D.L. Riddle. 1993. The daf-4 gene encodes a bone morphogenetic protein receptor controlling C. elegans dauer larva development. Nature **365**:644-649.

Evans, M.J. and M.H. Kaufman. 1981. Establishment in culture of pluripotential cells from mouse embryos. Nature **292**:154-156.

Farini, E. and C.B. Whitelaw. 1995. Ectopic expression of beta-lactoglobulin transgenes. Mol Gen Genet **246**:734-738.

Feinberg, A.P. and B. Vogelstein . 1983. A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal Biochem **132**:6-13.

Feldman, B., W. Poueymirou, V.E. Papaioannou, T.M. DeChiara, and M. Goldfarb. 1995. Requirement of FGF-4 for postimplantation mouse development. Science **267**:246-249.

Feng, X.H. and R. Derynck. 1997. A kinase subdomain of transforming growth factor-beta (TGF-beta) type I receptor determines the TGF-beta intracellular signaling specificity. EMBO J **16**:3912-3923.

Ferguson, E.L. and K.V. Anderson . 1992. Decapentaplegic acts as a morphogen to organize dorsal-ventral pattern in the Drosophila embryo. Cell **71**:451-461.

Fiering, S., E. Epner, K. Robinson, Y. Zhuang, A. Telling, M. Hu, D.I. Martin, T. Enver, T.J. Ley, and M. Groudine. 1995. Targeted deletion of 5'HS2 of the murine beta-globin LCR reveals that it is not essential for proper regulation of the beta-globin locus. Genes Dev **9**:2203-2213.

Franzen, P., H. Ichijo, and K. Miyazono. 1993. Different signals mediate transforming growth factor-beta 1-induced growth inhibition and extracellular matrix production in prostatic carcinoma cells. Exp Cell Res **207**:1-7.

Franzen, P., P. ten Dijke, H. Ichijo, H. Yamashita, P. Schulz, C.H. Heldin, and K. Miyazono. 1993. Cloning of a TGF beta type I receptor that forms a heteromeric complex with the TGF beta type II receptor. Cell **75**:681-692.

Franzen, P., C.H. Heldin, and K. Miyazono. 1995. The GS domain of the transforming growth factor-beta type I receptor is important in signal transduction. Biochem Biophys Res Commun **207**:682-689.

Fukada, T., T. Ohtani, Y. Yoshida, T. Shirogane, K. Nishida, K. Nakajima, M. Hibi, and T. Hirano. 1998. STAT3 orchestrates contradictory signals in cytokine-induced G1 to S cell-cycle transition. EMBO J **17**:6670-6677.

Gaye, P, Hue-Delahaie, D, Mercier, J.C, Soulier, S, Vilotte, J.L, Furet, J.P. 1986. Ovine beta-lactoglobulin messenger RNA: nucleotide sequence and mRNA levels during functional differentiation of the mammary gland. Biochimie **68** 1097-1107

Geng, Y. and R.A. Weinberg. 1993. Transforming growth factor beta effects on expression of G1 cyclins and cyclin-dependent protein kinases. Proc Natl Acad Sci U S A **90** :10315-10319.

Gerritsen, M.E., A.J. Williams, A.S. Neish, S. Moore, Y. Shi, and T. Collins. 1997. CREB-binding protein/p300 are transcriptional coactivators of p65. Proc Natl Acad Sci U S A **94**:2927-2932.

Gold, L.I. 1999. The role for transforming growth factor-beta (TGF-beta) in human cancer. Crit Rev Oncog **10**:303-360.

Goodnow, C.C. 1992. Transgenic mice and analysis of B-cell tolerance. Annu Rev Immunol **10**:489-518.

Gordon, J.W., G.A. Scangos, D.J. Plotkin, J.A. Barbosa, and F.H. Ruddle. 1980. Genetic transformation of mouse embryos by microinjection of purified DNA. Proc Natl Acad Sci U S A **77**:7380-7384.

Gordon, J.W. and F.H. Ruddle. 1981. Integration and stable germ line transmission of genes injected into mouse pronuclei. Science **214**:1244-1246.

Gossen, M. and H. Bujard. 1992. Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. Proc Natl Acad Sci U S A **89**:5547-5551.

Gossen, M., S. Freundlieb, G. Bender, G. Muller, W. Hillen, and H. Bujard. 1995. Transcriptional activation by tetracyclines in mammalian cells. Science **268**:1766-1769.

Goumans, M.J., D. Ward-van Costwaard, F. Wianny, P. Savatier, A. Zwijsen, and C. Mummery. 1998. Mouse embryonic stem cells with aberrant transforming growth factor beta signalling exhibit impaired differentiation in vitro and in vivo. Differentiation **63**:101-113.

Graff, J.M., A. Bansal, and D.A. Melton. 1996. Xenopus Mad proteins transduce distinct subsets of signals for the TGF beta superfamily. Cell **85**:479-487.

Grau, A.M., L. Zhang, W. Wang, S. Ruan, D.B. Evans, J.L. Abbruzzese, W. Zhang, and P.J. Chiao. 1997. Induction of p21waf1 expression and growth inhibition by transforming

growth factor beta involve the tumor suppressor gene DPC4 in human pancreatic adenocarcinoma cells. Cancer Res **57**:3929-3934.

Green, J.B. 1994. Roads to neuralness: embryonic neural induction as derepression of a default state. Cell **77**:317-320.

Grieder, N.C., D. Nellen, R. Burke, K. Basler, and M. Affolter. 1995. Schnurri is required for Drosophila Dpp signaling and encodes a zinc finger protein similar to the mammalian transcription factor PRDII-BF1. Cell **81**:791-800.

Gu, H., J.D. Marth, P.C. Orban, H. Mossmann, and K. Rajewsky. 1994. Deletion of a DNA polymerase beta gene segment in T cells using cell type-specific gene targeting [see comments]. Science **265**:103-106.

Hahn, S.A., A.T. Hoque, C.A. Moskaluk, L.T. da Costa, M. Schutte, E. Rozenblum, A.B. Seymour, C.L. Weinstein, C.J. Yeo, R.H. Hruban, and S.E. Kern. 1996. Homozygous deletion map at 18q21.1 in pancreatic cancer. Cancer Res **56**:490-494.

Hall, A.L., J. Franke, M. Faure, and R.H. Kessin. 1993. The role of the cyclic nucleotide phosphodiesterase of Dictyostelium discoideum during growth, aggregation, and morphogenesis: overexpression and localization studies with the separate promoters of the pde. Dev Biol **157**:73-84.

Hammer, R.E., R.D. Palmiter, and R.L. Brinster. 1984. Partial correction of murine hereditary growth disorder by germ-line incorporation of a new gene. Nature **311**:65-67.

Hanahan, D. 1989. Transgenic mice as probes into complex systems. Science **246**:1265-1275.

Hanai, J. and K. Miyazono. 2000. [PEBP2/CBF is a nuclear target of Smads in TGF-beta super family signaling]. Tanpakushitsu Kakusan Koso **45**:48-54.

Harbers, K., D. Jahner, and R. Jaenisch. 1981. Microinjection of cloned retroviral genomes into mouse zygotes: integration and expression in the animal. Nature **293**:540-542.

Harris, S., M. McClenaghan, J.P. Simons, S. Ali, and A.J. Clark. 1991. Developmental regulation of the sheep beta-lactoglobulin gene in the mammary gland of transgenic mice. Dev Genet **12**:299-307.

Hasty, P., J. Rivera-Perez, C. Chang, and A. Bradley. 1991. Target frequency and integration pattern for insertion and replacement vectors in embryonic stem cells. Mol Cell Biol **11**:4509-4517.

Hata, A., R.S. Lo, D. Wotton, G. Lagna, and J. Massague. 1997. Mutations increasing autoinhibition inactivate tumour suppressors Smad2 and Smad4 [see comments]. Nature **388**:82-87.

Hawley, S.H., K. Wunnenberg-Stapleton, C. Hashimoto, M.N. Laurent, T. Watabe, B.W. Blumberg, and K.W. Cho. 1995. Disruption of BMP signals in embryonic Xenopus ectoderm leads to direct neural induction. Genes Dev **9**:2923-2935.

Heberlein, U., T. Wolff, and G.M. Rubin. 1993. The TGF beta homolog dpp and the segment polarity gene hedgehog are required for propagation of a morphogenetic wave in the Drosophila retina. Cell **75**:913-926.

Heine, U., E.F. Munoz, K.C. Flanders, L.R. Ellingsworth, H.Y. Lam, N.L. Thompson, A.B. Roberts, and M.B. Sporn. 1987. Role of transforming growth factor-beta in the development of the mouse embryo. J Cell Biol **105**:2861-2876.

Heldin, C.H., K. Miyazono, and P. ten Dijke. 1997. TGF-beta signalling from cell membrane to nucleus through SMAD proteins. Nature **390**:465-471.

Hemmati-Brivanlou, A. and D. Melton. 1997. Vertebrate neural induction. Annu Rev Neurosci **20**:43-60.

Hennighausen, L., R. McKnight, T. Burdon, M. Baik, R.J. Wall, and G.H. Smith. 1994. Whey acidic protein extrinsically expressed from the mouse mammary tumor virus long terminal repeat results in hyperplasia of the coagulation gland epithelium and impaired mammary development. Cell Growth Differ **5**:607-613.

Hennighausen, L., R.J. Wall, U. Tillmann, M. Li, and P.A. Furth. 1995. Conditional gene expression in secretory tissues and skin of transgenic mice using the MMTV-LTR and the tetracycline responsive system. J Cell Biochem **59**:463-472.

Hennighausen, L. and G.W. Robinson. 1998. Think globally, act locally: the making of a mouse mammary gland. Genes Dev **12**:449-455.

Hocevar, B.A., T.L. Brown, and P.H. Howe. 1999. TGF-beta induces fibronectin synthesis through a c-Jun N-terminal kinase-dependent, Smad4-independent pathway. EMBO J **18**:1345-1356.

Hoess, R.H., M. Ziese, and N. Sternberg. 1982. P1 site-specific recombination: nucleotide sequence of the recombining sites. Proc Natl Acad Sci U S A **79**:3398-3402.

Hoess, R.H. and K. Abremski. 1985. Mechanism of strand cleavage and exchange in the Cre-lox site-specific recombination system. J Mol Biol **181**:351-362.

Hogan, Beddington, F. Costantini, and Lacy. 1994. Manipulating the Mouse Embryo. Cold Spring Harbor Laboratory Press,

Hogan, B.L., M. Blessing, G.E. Winnier, N. Suzuki, and C.M. Jones. 1994. Growth factors in development: the role of TGF-beta related polypeptide signalling molecules in embryogenesis. Dev Suppl 53-60.

Hogan, B.L. 1994. Developmental signalling. Sorting out the signals. Curr Biol **4**:1122-1124.

Hogan, B.L. 1996. Bone morphogenetic proteins: multifunctional regulators of vertebrate development. Genes Dev **10**:1580-1594.

Holley, S.A., J.L. Neul, L. Attisano, J.L. Wrana, Y. Sasai, M.B. O'Connor, E.M. De Robertis, and E.L. Ferguson. 1996. The Xenopus dorsalizing factor noggin ventralizes Drosophila embryos by preventing DPP from activating its receptor. Cell **86**:607-617.

Holliday R. 1964. A mechanism for gene conversion in fungi. Genetic Research **5**, 282-304.

Hoodless, P.A., T. Haerry, S. Abdollah, M. Stapleton, M.B. O'Connor, L. Attisano, and J.L. Wrana. 1996. MADR1, a MAD-related protein that functions in BMP2 signaling pathways. Cell **85**:489-500.

Hooper M, Hardy K, Handyside A, Hunter S, and Monk M. 1987. HPRT-deficient (Lesch-Nyhan) mouse embryos derived from germline colonization by cultured cells. Nature **326**(6110), 292-295

Hooper, M.L. 1992. EMBRYONAL STEM CELLS - Introducing Planned Changes into the Animal Germline. Harwood academic publishers,

Howell, M., F. Itoh, C.E. Pierreux, S. Valgeirsdottir, S. Itoh, P. ten Dijke, and C.S. Hill. 1999. Xenopus Smad4beta is the co-Smad component of developmentally regulated transcription factor complexes responsible for induction of early mesodermal genes. Dev Biol **214**:354-369.

Hua, X., X. Liu, D.O. Ansari, and H.F. Lodish. 1998. Synergistic cooperation of TFE3 and smad proteins in TGF-beta-induced transcription of the plasminogen activator inhibitor-1 gene. Genes Dev **12**:3084-3095.

Hynes, R.O. 1992. Integrins: versatility, modulation, and signaling in cell adhesion. Cell **69**:11-25.

Imbeaud, S., E. Faure, I. Lamarre, M.G. Mattei, N. di Clemente, R. Tizard, D. Carre-Eusebe, C. Belville, L. Tragethon, and C. Tonkin. 1995. Insensitivity to anti-mullerian hormone due to a mutation in the human anti-mullerian hormone receptor. Nat Genet **11**:382-388.

Irish, V.F. and W.M. Gelbart. 1987. The decapentaplegic gene is required for dorsal-ventral patterning of the Drosophila embryo. Genes Dev **1**:868-879.

Ishisaki, A., K. Yamato, A. Nakao, K. Nonaka, M. Ohguchi, P. ten Dijke, and T. Nishihara. 1998. Smad7 is an activin-inducible inhibitor of activin-induced growth arrest and apoptosis in mouse B cells. J Biol Chem **273**:24293-24296.

Jaenisch, R. and B. Mintz. 1974. Simian virus 40 DNA sequences in DNA of healthy adult mice derived from preimplantation blastocysts injected with viral DNA. Proc Natl Acad Sci U S A **71**:1250-1254.

Jaenisch, R. 1976. Germ line integration and Mendelian transmission of the exogenous Moloney leukemia virus. Proc Natl Acad Sci U S A **73**:1260-1264.

Jaenisch, R. 1988. Transgenic animals. Science **240**:1468-1474.

Janknecht, R., N.J. Wells, and T. Hunter. 1998. TGF-beta-stimulated cooperation of smad proteins with the coactivators CBP/p300. Genes Dev **12**:2114-2119.

Jhappan, C., C. Stahle, R.N. Harkins, N. Fausto, G.H. Smith, and G.T. Merlino. 1990. TGF alpha overexpression in transgenic mice induces liver neoplasia and abnormal development of the mammary gland and pancreas. Cell **61**:1137-1146.

Jhappan, C., A.G. Geiser, E.C. Kordon, D. Bagheri, L. Hennighausen, A.B. Roberts, G.H. Smith, and G. Merlino. 1993. Targeting expression of a transforming growth factor beta 1 transgene to the pregnant mammary gland inhibits alveolar development and lactation. EMBO J **12**:1835-1845.

Johnson, K., H. Kirkpatrick, A. Comer, F.M. Hoffmann, and A. Laughon. 1999. Interaction of Smad complexes with tripartite DNA-binding sites. J Biol Chem **274**:20709-20716.

Jones, C.M., N. Armes, and J.C. Smith. 1996. Signalling by TGF-beta family members: short-range effects of Xnr-2 and BMP-4 contrast with the long-range effects of activin. Curr Biol **6**:1468-1475.

Kaartinen, V., J.W. Voncken, C. Shuler, D. Warburton, D. Bu, N. Heisterkamp, and J. Groffen. 1995. Abnormal lung development and cleft palate in mice lacking TGF-beta 3 indicates defects of epithelial-mesenchymal interaction. Nat Genet **11**:415-421.

Kanegae, Y., G. Lee, Y. Sato, M. Tanaka, M. Nakai, T. Sakaki, S. Sugano, and I. Saito. 1995. Efficient gene activation in mammalian cells by using recombinant adenovirus expressing site-specific Cre recombinase. Nucleic Acids Res **23**:3816-3821.

Kaplan, M.H., C. Daniel, U. Schindler, and M.J. Grusby. 1998. Stat proteins control lymphocyte proliferation by regulating p27Kip1 expression. Mol Cell Biol **18**:1996-2003.

Katsuki, M., M. Sato, M. Kimura, M. Yokoyama, K. Kobayashi, and T. Nomura. 1988. Conversion of normal behavior to shiverer by myelin basic protein antisense cDNA in transgenic mice. Science **241**:593-595.

Kawabata, M., T. Imamura, K. Miyazono, M.E. Engel, and H.L. Moses. 1995. Interaction of the transforming growth factor-beta type I receptor with farnesyl-protein transferase-alpha. J Biol Chem **270**:29628-29631.

Kawase, E., H. Suemori, N. Takahashi, K. Okazaki, K. Hashimoto, and N. Nakatsuji. 1994. Strain difference in establishment of mouse embryonic stem (ES) cell lines. Int J Dev Biol **38**:385-390.

Kellendonk, C., C. Opherk, K. Anlag, G. Schutz, and F. Tronche. 2000. Hepatocyte-specific expression of Cre recombinase. Genesis **26**:151-153.

Kennedy, M.K., D.S. Torrance, K.S. Picha, and K.M. Mohler. 1992. Analysis of cytokine mRNA expression in the central nervous system of mice with experimental autoimmune encephalomyelitis reveals that IL-10 mRNA expression correlates with recovery. J Immunol **149**:2496-2505.

Kessel, M. and P. Gruss. 1990. Murine developmental control genes. Science **249**:374-379.

Khoury, S.J., W.W. Hancock, and H.L. Weiner. 1992. Oral tolerance to myelin basic protein and natural recovery from experimental autoimmune encephalomyelitis are associated with downregulation of inflammatory cytokines and differential upregulation of transforming growth factor beta, interleukin 4, and prostaglandin E expression in the brain. J Exp Med **176**:1355-1364.

Kim, J., K. Johnson, H.J. Chen, S. Carroll, and A. Laughon. 1997. Drosophila Mad binds to DNA and directly mediates activation of vestigial by Decapentaplegic. Nature **388**:304-308.

Kim, J.W., H.S. Kim, I.K. Kim, M.R. Kim, E.Y. Cho, H.K. Kim, J.M. Lee, and S.E. Namkoong. 1998. Transforming growth factor-beta 1 induces apoptosis through down-regulation of c-myc gene and overexpression of p27Kip1 protein in cervical carcinoma. Gynecol Oncol **69**:230-236.

Kingsley, D.M., A.E. Bland, J.M. Grubber, P.C. Marker, L.B. Russell, N.G. Copeland, and N.A. Jenkins. 1992. The mouse short ear skeletal morphogenesis locus is associated

with defects in a bone morphogenetic member of the TGF beta superfamily. Cell **71**:399-410.

Kocarek, T.A., R.C. Zangar, and R.F. Novak. 2000. Post-transcriptional regulation of rat CYP2E1 expression: role of CYP2E1 mRNA untranslated regions in control of translational efficiency and message stability. Arch Biochem Biophys **376**:180-190.

Korach, K.S., J.F. Couse, S.W. Curtis, T.F. Washburn, J. Lindzey, K.S. Kimbro, E.M. Eddy, S. Migliaccio, S.M. Snedeker, D.B. Lubahn, D.W. Schomberg, and E.P. Smith. 1996. Estrogen receptor gene disruption: molecular characterization and experimental and clinical phenotypes. Recent Prog Horm Res **51**:159-158.

Kordon, E.C., R.A. McKnight, C. Jhappan, L. Hennighausen, G. Merlino, and G.H. Smith. 1995. Ectopic TGF beta 1 expression in the secretory mammary epithelium induces early senescence of the epithelial stem cell population. Dev Biol **168**:47-61.

Kozak, M. 1986. Influences of mRNA secondary structure on initiation by eukaryotic ribosomes. Proc Natl Acad Sci U S A **83**:2850-2854.

Kretzschmar, M., J. Doody, and J. Massague. 1997. Opposing BMP and EGF signalling pathways converge on the TGF-beta family mediator Smad1. Nature **389**:618-622.

Kretzschmar, M., F. Liu, A. Hata, J. Doody, and J. Massague. 1997. The TGF-beta family mediator Smad1 is phosphorylated directly and activated functionally by the BMP receptor kinase. Genes Dev **11**:984-995.

Kretzschmar, M. and J. Massague. 1998. SMADs: mediators and regulators of TGF-beta signaling. Curr Opin Genet Dev **8**:103-111.

Kretzschmar, M., J. Doody, I. Timokhina, and J. Massague. 1999. A mechanism of repression of TGFbeta/ Smad signaling by oncogenic Ras. Genes Dev **13**:804-816.

Krishna, S., L.L. Maduzia, and R.W. Padgett. 1999. Specificity of TGFbeta signaling is conferred by distinct type I receptors and their associated SMAD proteins in Caenorhabditis elegans. Development **126**:251-260.

Kulkarni, A.B. and S. Karlsson. 1993. Transforming growth factor-beta 1 knockout mice. A mutation in one cytokine gene causes a dramatic inflammatory disease. Am J Pathol **143**:3-9.

Kulkarni, A.B., C.G. Huh, D. Becker, A. Geiser, M. Lyght, K.C. Flanders, A.B. Roberts, M.B. Sporn, J.M. Ward, and S. Karlsson. 1993. Transforming growth factor beta 1 null mutation in mice causes excessive inflammatory response and early death. Proc Natl Acad Sci U S A **90**:770-774.

Kumar, A., M. Commane, T.W. Flickinger, C.M. Horvath, and G.R. Stark. 1997. Defective TNF-alpha-induced apoptosis in STAT1-null cells due to low constitutive levels of caspases [see comments]. Science **278**:1630-1632.

Kumar, T.R., Y. Wang, and M.M. Matzuk. 1996. Gonadotropins are essential modifier factors for gonadal tumor development in inhibin-deficient mice. Endocrinology **137**:4210-4216.

Kurokowa, M., K. Lynch, and D.K. Podolsky. 1987. Effects of growth factors on an intestinal epithelial cell line: transforming growth factor beta inhibits proliferation and stimulates differentiation. Biochem Biophys Res Commun **142**:775-782.

Labbe, E., C. Silvestri, P.A. Hoodless, J.L. Wrana, and L. Attisano. 1998. Smad2 and Smad3 positively and negatively regulate TGF beta-dependent transcription through the forkhead DNA-binding protein FAST2. Mol Cell **2**:109-120.

Lagna, G., A. Hata, A. Hemmati-Brivanlou, and J. Massague. 1996. Partnership between DPC4 and SMAD proteins in TGF-beta signalling pathways. Nature **383**:832-836.

Laiho, M., J.A. DeCaprio, J.W. Ludlow, D.M. Livingston, and J. Massague. 1990. Growth inhibition by TGF-beta linked to suppression of retinoblastoma protein phosphorylation. *Cell* **62**:175-185.

Laiho, M., M.B. Weis, and J. Massague. 1990. Concomitant loss of transforming growth factor (TGF)-beta receptor types I and II in TGF-beta-resistant cell mutants implicates both receptor types in signal transduction. *J Biol Chem* **265**:18518-18524.

Lakso, M., B. Sauer, B.J. Mosinger, E.J. Lee, R.W. Manning, S.H. Yu, K.L. Mulder, and H. Westphal. 1992. Targeted oncogene activation by site-specific recombination in transgenic mice. Proc Natl Acad Sci U S A **89**:6232-6236.

Letsou, A., K. Arora, J.L. Wrana, K. Simin, V. Twombly, J. Jamal, K. Staehling-Hampton, F.M. Hoffmann, W.M. Gelbart, and J. Massague. 1995. Drosophila Dpp signaling is mediated by the punt gene product: a dual ligand-binding type II receptor of the TGF beta receptor family. Cell **80**:899-908.

Letterio, J.J., A.G. Geiser, A.B. Kulkarni, N.S. Roche, M.B. Sporn, and A.B. Roberts. 1994. Maternal rescue of transforming growth factor-beta 1 null mice. Science **264** :1936-1938.

Lewandoski, M., K.M. Wassarman, and G.R. Martin. 1997. Zp3-cre, a transgenic mouse line for the activation or inactivation of loxP-flanked target genes specifically in the female germ line. Curr Biol **7**:148-151.

Li, X.Y., v.H. Hooft, R. Mantovani, C. Benoist, and D. Mathis. 1992. Intron-exon organization of the NF-Y genes. Tissue-specific splicing modifies an activation domain. J Biol Chem **267**:8984-8990.

Liberati, N.T., M.B. Datto, J.P. Frederick, X. Shen, C. Wong, E.M. Rougier-Chapman, and X.F. Wang 1999. Smads bind directly to the Jun family of AP-1 transcription factors. Proc Natl Acad Sci U S A **96**:4844-4849.

Liu, B., C.L. Dou, L. Prabhu, and E. Lai. 1999. FAST-2 is a mammalian winged-helix protein which mediates transforming growth factor beta signals. Mol Cell Biol **19**:424-430.

Liu, F., F. Ventura, J. Doody, and J. Massague. 1995. Human type II receptor for bone morphogenic proteins (BMPs): extension of the two-kinase receptor model to the BMPs. Mol Cell Biol **15**:3479-3486.

Liu, F., C. Pouponnot, and J. Massague. 1997. Dual role of the Smad4/DPC4 tumor suppressor in TGFbeta-inducible transcriptional complexes. Genes Dev **11**:3157-3167.

Liu, X., G.W. Robinson, K.U. Wagner, L. Garrett, A. Wynshaw-Boris, and L. Hennighausen. 1997. Stat5a is mandatory for adult mammary gland development and lactogenesis. Genes Dev **11**:179-186.

Lo, R.S., Y.G. Chen, Y. Shi, N.P. Pavletich, and J. Massague. 1998. The L3 loop: a structural motif determining specific interactions between SMAD proteins and TGF-beta receptors. EMBO J **17**:996-1005.

Ludwig, T., D.L. Chapman, V.E. Papaioannou, and A. Efstratiadis. 1997. Targeted mutations of breast cancer susceptibility gene homologs in mice: lethal phenotypes of Brca1, Brca2, Brca1/Brca2, Brca1/p53, and Brca2/p53 nullizygous embryos. Genes Dev **11**:1226-1241.

Luo, K. and H.F. Lodish. 1996. Signaling by chimeric erythropoietin-TGF-beta receptors: homodimerization of the cytoplasmic domain of the type I TGF-beta receptor and heterodimerization with the type II receptor are both required for intracellular signal transduction. EMBO J **15**:4485-4496.

Luo, K., S.L. Stroschein, W. Wang, D. Chen, E. Martens, S. Zhou, and Q. Zhou. 1999. The Ski oncoprotein interacts with the Smad proteins to repress TGFbeta signaling. Genes Dev **13**:2196-2206.

Lustig, K.D., K.L. Kroll, E.E. Sun, and M.W. Kirschner. 1996. Expression cloning of a Xenopus T-related gene (Xombi) involved in mesodermal patterning and blastopore lip formation. Development **122**:4001-4012.

Lyon MF. 1996. *Genetic variants and strains of the laboratory mouse, edn 3.* Oxford: Oxford University Press,

Lyons, R.M., L.E. Gentry, A.F. Purchio, and H.L. Moses. 1990. Mechanism of activation of latent recombinant transforming growth factor beta 1 by plasmin. J Cell Biol **110**:1361-1367.

Ma, C., Y. Zhou, P.A. Beachy, and K. Moses. 1993. The segment polarity gene hedgehog is required for progression of the morphogenetic furrow in the developing Drosophila eye. Cell **75**:927-938.

Macias-Silva, M., S. Abdollah, P.A. Hoodless, R. Pirone, L. Attisano, and J.L. Wrana. 1996. MADR2 is a substrate of the TGFbeta receptor and its phosphorylation is required for nuclear accumulation and signaling. Cell **87**:1215-1224.

Maeda, H. and A. Shiraishi. 1996. TGF-beta contributes to the shift toward Th2-type responses through direct and IL-10-mediated pathways in tumor-bearing mice. J Immunol **156**:73-78.

Magin, T.M., C. McEwan, M. Milne, A.M. Pow, J. Selfridge, and D.W. Melton. 1992. A position- and orientation-dependent element in the first intron is required for expression of the mouse hprt gene in embryonic stem cells. Gene **122**:289-296.

Magin, T.M., J. McWhir, and D.W. Melton. 1992. A new mouse embryonic stem cell line with good germ line contribution and gene targeting frequency. Nucleic Acids Res **20**:3795-3796.

Manak, J.R., L.D. Mathies, and M.P. Scott. 1994. Regulation of a decapentaplegic midgut enhancer by homeotic proteins. Development **120**:3605-3619.

Mansour, S.L., K.R. Thomas, and M.R. Capecchi. 1988. Disruption of the proto-oncogene int-2 in mouse embryo-derived stem cells: a general strategy for targeting mutations to non-selectable genes. Nature **336**:348-352.

Mansour, S.L. and G.R. Martin. 1988. Four classes of mRNA are expressed from the mouse int-2 gene, a member of the FGF gene family. EMBO J **7**:2035-2041.

Martin, G.R. 1981. Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. Proc Natl Acad Sci U S A **78**:7634-7638.

Martin, P., M.C. Dickson, F.A. Millan, and R.J. Akhurst. 1993. Rapid induction and clearance of TGF beta 1 is an early response to wounding in the mouse embryo. Dev Genet **14**:225-238.

Massague, J. 1990. The transforming growth factor-beta family. Annu Rev Cell Biol **6**:597-641.

Massague, J. and F. Weis-Garcia. 1996. Serine/threonine kinase receptors: mediators of transforming growth factor beta family signals. Cancer Surv **27**:41-64.

Massague, J. 1998. TGF-beta signal transduction. Annu Rev Biochem **67**:753-791.

Massague, J. and D. Wotton. 2000. Transcriptional control by the TGF-beta/Smad signaling system. EMBO J **19**:1745-1754.

Masucci, J.D. and F.M. Hoffmann. 1993. Identification of two regions from the Drosophila decapentaplegic gene required for embryonic midgut development and larval viability. Dev Biol **159**:276-287.

Mathews, L.S. and W.W. Vale. 1991. Expression cloning of an activin receptor, a predicted transmembrane serine kinase. Cell **65**:973-982.

Matzuk, M.M., M.J. Finegold, J.G. Su, A.J. Hsueh, and A. Bradley. 1992. Alpha-inhibin is a tumour-suppressor gene with gonadal specificity in mice. Nature **360**:313-319.

Matzuk, M.M., M.J. Finegold, Y. Mishina, A. Bradley, and R.R. Behringer. 1995. Synergistic effects of inhibins and mullerian-inhibiting substance on testicular tumorigenesis. Mol Endocrinol **9**:1337-1345.

Mazars, P., N. Barboule, V. Baldin, S. Vidal, B. Ducommun, and A. Valette. 1995. Effects of TGF-beta 1 (transforming growth factor-beta 1) on the cell cycle regulation of human breast adenocarcinoma (MCF-7) cells. FEBS Lett **362**:295-300.

Medina, D., C.J. Oborn, F.S. Kittrell, and R.L. Ullrich. 1986. Properties of mouse mammary epithelial cell lines characterized by in vivo transplantation and in vitro immunocytochemical methods. J Natl Cancer Inst **76**:1143-1156.

Miech, R.P., R.E.J. Parks, J.H.J. Anderson, and A.C. Sartorelli. 1967. An hypothesis on the mechanism of action of 6-thioguanine. Biochem Pharmacol **16**:2222-2227.

Mishina, Y., A. Suzuki, N. Ueno, and R.R. Behringer. 1995. Bmpr encodes a type I bone morphogenetic protein receptor that is essential for gastrulation during mouse embryogenesis. Genes Dev **9**:3027-3037.

Miyazono, K. and C.H. Heldin. 1993. The mechanism of action of transforming growth factor-beta. Gastroenterol Jpn **28 Suppl 4**:81-87.

Miyazono, K., H. Ichijo, and C.H. Heldin. 1993. Transforming growth factor-beta: latent forms, binding proteins and receptors. Growth Factors **8**:11-22.

Mohanty, S., K.A. Jermyn, A. Early, T. Kawata, L. Aubry, A. Ceccarelli, P. Schaap, J.G. Williams, and R.A. Firtel. 1999. Evidence that the Dictyostelium Dd-STATa protein is a repressor that regulates commitment to stalk cell differentiation and is also required for efficient chemotaxis. Development **126**:3391-3405.

Moreau, J.F., D.D. Donaldson, F. Bennett, J. Witek-Giannotti, S.C. Clark, and G.G. Wong. 1988. Leukaemia inhibitory factor is identical to the myeloid growth factor human interleukin for DA cells. Nature **336**:690-692.

Morita, K., K.L. Chow, and N. Ueno. 1999. Regulation of body length and male tail ray pattern formation of Caenorhabditis elegans by a member of TGF-beta family. Development **126**:1337-1347.

Morita, K., K.L. Chow, and N. Ueno. 1999. Regulation of body length and male tail ray pattern formation of Caenorhabditis elegans by a member of TGF-beta family. Development **126**:1337-1347.

Motyl, T., K. Grzelkowska, W. Zimowska, J. Skierski, P. Wareski, T. Ploszaj, and L. Trzeciak. 1998. Expression of bcl-2 and bax in TGF-beta 1-induced apoptosis of L1210 leukemic cells. Eur J Cell Biol **75**:367-374.

Muenke, M. 1995. Finding genes involved in human developmental disorders. Curr Opin Genet Dev **5**:354-361.

Muller, W.J., E. Sinn, P.K. Pattengale, R. Wallace, and P. Leder. 1988. Single-step induction of mammary adenocarcinoma in transgenic mice bearing the activated c-neu oncogene. Cell **54**:105-115.

Murphy, M.S. 1998. Growth factors and the gastrointestinal tract. Nutrition **14**:771-774.

Nachtigal, M.W. and H.A. Ingraham. 1996. Bioactivation of Mullerian inhibiting substance during gonadal development by a kex2/subtilisin-like endoprotease. Proc Natl Acad Sci U S A **93**:7711-7716.

Nakajima, K., Y. Yamanaka, K. Nakae, H. Kojima, M. Ichiba, N. Kiuchi, T. Kitaoka, T. Fukada, M. Hibi, and T. Hirano. 1996. A central role for Stat3 in IL-6-induced regulation of growth and differentiation in M1 leukemia cells. EMBO J **15**:3651-3658.

Nakashima, K., M. Yanagisawa, H. Arakawa, N. Kimura, T. Hisatsune, M. Kawabata, K. Miyazono, and T. Taga. 1999. Synergistic signaling in fetal brain by STAT3-Smad1 complex bridged by p300 [see comments]. Science **284**:479-482.

Nandan, D. and N.E. Reiner. 1997. TGF-beta attenuates the class II transactivator and reveals an accessory pathway of IFN-gamma action. J Immunol **158**:1095-1101.

Neeper, M., R. Lowe, S. Galuska, K.J. Hofmann, R.G. Smith, and A. Elbrecht. 1996. Molecular cloning of an avian anti-Mullerian hormone homologue. Gene **176**:203-209.

Nellen, D., M. Affolter, and K. Basler. 1994. Receptor serine/threonine kinases implicated in the control of Drosophila body pattern by decapentaplegic. Cell **78**:225-237.

Nemetz, C. and G.M. Hocke. 1998. Transcription factor Stat5 is an early marker of differentiation of murine embryonic stem cells. Differentiation **62**:213-220.

Nishihara, A., J.I. Hanai, N. Okamoto, J. Yanagisawa, S. Kato, K. Miyazono, and M. Kawabata. 1998. Role of p300, a transcriptional coactivator, in signalling of TGF-beta. Genes Cells **3**:613-623.

Nishimori, K., L.J. Young, Q. Guo, Z. Wang, T.R. Insel, and M.M. Matzuk. 1996. Oxytocin is required for nursing but is not essential for parturition or reproductive behavior. Proc Natl Acad Sci U S A **93**:11699-11704.

Ogryzko, V.V., R.L. Schiltz, V. Russanova, B.H. Howard, and Y. Nakatani. 1996. The transcriptional coactivators p300 and CBP are histone acetyltransferases. Cell **87**:953-959.

Olson, E.N., H.H. Arnold, P.W. Rigby, and B.J. Wold. 1996. Know your neighbors: three phenotypes in null mutants of the myogenic bHLH gene MRF4. Cell **85**:1-4.

Orban, P.C., D. Chui, and J.D. Marth. 1992. Tissue- and site-specific DNA recombination in transgenic mice. Proc Natl Acad Sci U S A **89**:6861-6865.

Ormerod, E.J. and P.S. Rudland. 1985. Isolation and differentiation of cloned epithelial cell lines from normal rat mammary glands. In Vitro Cell Dev Biol **21**:143-153.

Ostrom, L., M.J. Tang, P. Gruss, and G.R. Dressler. 2000. Reduced Pax2 gene dosage increases apoptosis and slows the progression of renal cystic disease. Dev Biol **219**:250-258.

Palmiter, R.D., R.L. Brinster, R.E. Hammer, M.E. Trumbauer, M.G. Rosenfeld, N.C. Birnberg, and R.M. Evans. 1982. Dramatic growth of mice that develop from eggs microinjected with metallothionein-growth hormone fusion genes. Nature **300**:611-615.

Palmiter, R.D., T.M. Wilkie, H.Y. Chen, and R.L. Brinster. 1984. Transmission distortion and mosaicism in an unusual transgenic mouse pedigree. Cell **36**:869-877.

Palmiter, R.D. and R.L. Brinster . 1986. Germ-line transformation of mice. Annu Rev Genet **20**:465-499.

Panganiban, G.E., R. Reuter, M.P. Scott, and F.M. Hoffmann. 1990. A Drosophila growth factor homolog, decapentaplegic, regulates homeotic gene expression within and across germ layers during midgut morphogenesis. Development **110**:1041-1050.

Pardoux, C., C. Asselin-Paturel, J. Chehimi, F. Gay, F. Mami-Chouaib, and S. Chouaib. 1997. Functional interaction between TGF-beta and IL-12 in human primary allogeneic cytotoxicity and proliferative response. J Immunol **158**:136-143.

Pearson, K.L., T. Hunter, and R. Janknecht. 1999. Activation of Smad1-mediated transcription by p300/CBP. Biochim Biophys Acta **1489**:354-364.

Pease, S. and R.L. Williams. 1990. Formation of germ-line chimeras from embryonic stem cells maintained with recombinant leukemia inhibitory factor. Exp Cell Res **190**:209-211.

Penton, A., Y. Chen, K. Staehling-Hampton, J.L. Wrana, L. Attisano, J. Szidonya, J.A. Cassill, J. Massague, and F.M. Hoffmann. 1994. Identification of two bone morphogenetic protein type I receptors in Drosophila and evidence that Brk25D is a decapentaplegic receptor. Cell **78**:239-250.

Piccolo, S., Y. Sasai, B. Lu, and E.M. De Robertis. 1996. Dorsoventral patterning in Xenopus: inhibition of ventral signals by direct binding of chordin to BMP-4. Cell **86**:589-598.

Pierce, D.F.J., M.D. Johnson, Y. Matsui, S.D. Robinson, L.I. Gold, A.F. Purchio, C.W. Daniel, B.L. Hogan, and H.L. Moses. 1993. Inhibition of mammary duct development but not alveolar outgrowth during pregnancy in transgenic mice expressing active TGF-beta 1. Genes Dev **7**:2308-2317.

Pollard, J.W. and L. Hennighausen. 1994. Colony stimulating factor 1 is required for mammary gland development during pregnancy. Proc Natl Acad Sci U S A **91**:9312-9316.

Potocnik, A.J, Brakebusch, C, Fassler, R. 2000. Fetal and adult hematopoietic stem cells require beta1 integrin function for colonizing fetal liver, spleen, and bone marrow. Immunity **12**: 653-663

Pouliot, F. and C. Labrie. 1999. Expression profile of agonistic Smads in human breast cancer cells: absence of regulation by estrogens. Int J Cancer **81**:98-103.

Proetzel, G., S.A. Pawlowski, M.V. Wiles, M. Yin, G.P. Boivin, P.N. Howles, J. Ding, M.W. Ferguson, and T. Doetschman. 1995. Transforming growth factor-beta 3 is required for secondary palate fusion. Nat Genet **11**:409-414.

Raftery, L.A., V. Twombly, K. Wharton, and W.M. Gelbart. 1995. Genetic screens to identify elements of the decapentaplegic signaling pathway in Drosophila. Genetics **139**:241-254.

Raftery, L.A. and R.G. Wisotzkey . 1996. Characterization of Medea, a gene required for maximal function of the Drosophila BMP homolog Decapentaplegic. Ann N Y Acad Sci **785**:318-320.

Raine, C.S. 1984. Biology of disease. Analysis of autoimmune demyelination: its impact upon multiple sclerosis. Lab Invest **50**:608-635.

Ray, M.K., S.P. Fagan, S. Moldovan, F.J. DeMayo, and F.C. Brunicardi. 1999. Development of a transgenic mouse model using rat insulin promoter to drive the expression of CRE recombinase in a tissue-specific manner. Int J Pancreatol **25**:157-163.

Reiss, M. and M.H. Barcellos-Hoff. 1997. Transforming growth factor-beta in breast cancer: a working hypothesis. Breast Cancer Res Treat **45**:81-95.

Renucci, A., V. Lemarchandel, and F. Rosa. 1996. An activated form of type I serine/threonine kinase receptor TARAM-A reveals a specific signalling pathway involved in fish head organiser formation. Development **122**:3735-3743.

Riggins, G.J., S. Thiagalingam, E. Rozenblum, C.L. Weinstein, S.E. Kern, S.R. Hamilton, J.K. Willson, S.D. Markowitz, K.W. Kinzler, and B. Vogelstein. 1996. Mad-related genes in the human. Nat Genet **13**:347-349.

Riggins, G.J., K.W. Kinzler, B. Vogelstein, and S. Thiagalingam. 1997. Frequency of Smad gene mutations in human cancers. Cancer Res **57**:2578-2580.

Roberts, A.B., K.C. Flanders, U.I. Heine, S. Jakowlew, P. Kondaiah, S.J. Kim, and M.B. Sporn. 1990. Transforming growth factor-beta: multifunctional regulator of differentiation and development. Philos Trans R Soc Lond B Biol Sci **327**:145-154.

Roberts, A.B., P. Kondaiah, F. Rosa, S. Watanabe, P. Good, D. Danielpour, N.S. Roche, M.L. Rebbert, I.B. Dawid, and M.B. Sporn. 1990. Mesoderm induction in Xenopus laevis distinguishes between the various TGF-beta isoforms. Growth Factors **3**:277-286.

Roberts, A.B., U.I. Heine, K.C. Flanders, and M.B. Sporn. 1990. Transforming growth factor-beta. Major role in regulation of extracellular matrix. Ann N Y Acad Sci **580** :225-232.

Robinson, S.D., G.B. Silberstein, A.B. Roberts, K.C. Flanders, and C.W. Daniel. 1991. Regulated expression and growth inhibitory effects of transforming growth factor-beta isoforms in mouse mammary gland development. Development **113**:867-878.

Rocken, M., J. Urban, and E.M. Shevach. 1994. Antigen-specific activation, tolerization, and reactivation of the interleukin 4 pathway in vivo. J Exp Med **179**:1885-1893.

Rohlmann, A., M. Gotthardt, T.E. Willnow, R.E. Hammer, and J. Herz. 1996. Sustained somatic gene inactivation by viral transfer of Cre recombinase [see comments]. Nat Biotechnol **14**:1562-1565.

Rosfjord, E.C. and R.B. Dickson. 1999. Growth factors, apoptosis, and survival of mammary epithelial cells. J Mammary Gland Biol Neoplasia **4**:229-237.

Rotello, R.J., R.C. Lieberman, A.F. Purchio, and L.E. Gerschenson. 1991. Coordinated regulation of apoptosis and cell proliferation by transforming growth factor beta 1 in cultured uterine epithelial cells. Proc Natl Acad Sci U S A **88**:3412-3415.

Ruberte, E., T. Marty, D. Nellen, M. Affolter, and K. Basler. 1995. An absolute requirement for both the type II and type I receptors, punt and thick veins, for dpp signaling in vivo. Cell **80**:889-897.

Ruegemer, J.J., S.N. Ho, J.A. Augustine, J.W. Schlager, M.P. Bell, D.J. McKean, and R.T. Abraham. 1990. Regulatory effects of transforming growth factor-beta on IL-2- and IL-4-dependent T cell-cycle progression. J Immunol **144**:1767-1776.

Sakakura, T. 1991. New aspects of stroma-parenchyma relations in mammary gland differentiation. Int Rev Cytol **125**:165-202.

Sambrook, Fritsch, and Maniatis. 1989. Molecular Cloning - A Laboratory Manual. Cold Spring Harbor Laboratory Press.

Sano, Y., J. Harada, S. Tashiro, R. Gotoh-Mandeville, T. Maekawa, and S. Ishii. 1999. ATF-2 is a common nuclear target of Smad and TAK1 pathways in transforming growth factor-beta signaling. J Biol Chem **274**:8949-8957.

Santambrogio, L., G.M. Hochwald, B. Saxena, C.H. Leu, J.E. Martz, J.A. Carlino, N.H. Ruddle, M.A. Palladino, L.I. Gold, and G.J. Thorbecke. 1993. Studies on the mechanisms by which transforming growth factor-beta (TGF-beta) protects against allergic encephalomyelitis. Antagonism between TGF-beta and tumor necrosis factor. J Immunol **151**:1116-1127.

Sauer, B. and N. Henderson. 1988. Site-specific DNA recombination in mammalian cells by the Cre recombinase of bacteriophage P1. Proc Natl Acad Sci U S A **85**:5166-5170.

Sauer, B. and N. Henderson. 1989. Cre-stimulated recombination at loxP-containing DNA sequences placed into the mammalian genome. Nucleic Acids Res **17**:147-161.

Sauer, B. 1996. Multiplex Cre/lox recombination permits selective site-specific DNA targeting to both a natural and an engineered site in the yeast genome. Nucleic Acids Res **24**:4608-4613.

Schindler, C. and J.E.J. Darnell . 1995. Transcriptional responses to polypeptide ligands: the JAK-STAT pathway. Annu Rev Biochem **64**:621-651.

Schluesener, H.J. and O. Lider. 1989. Transforming growth factors beta 1 and beta 2: cytokines with identical immunosuppressive effects and a potential role in the regulation of autoimmune T cell function. J Neuroimmunol **24**:249-258.

Schutte, M., R.H. Hruban, L. Hedrick, K.R. Cho, G.M. Nadasdy, C.L. Weinstein, G.S. Bova, W.B. Isaacs, P. Cairns, H. Nawroz, D. Sidransky, R.A.J. Casero, P.S. Meltzer, S.A. Hahn, and S.E. Kern. 1996. DPC4 gene in various tumor types. Cancer Res **56**:2527-2530.

Schwartzberg, P.L., S.P. Goff, and E.J. Robertson. 1989. Germ-line transmission of a c-abl mutation produced by targeted gene disruption in ES cells. Science **246**:799-803.

Segev, N. and G. Cohen. 1981. Control of circularization of bacteriophage P1 DNA in Escherichia coli. Virology **114**:333-342.

Sekelsky, J.J., S.J. Newfeld, L.A. Raftery, E.H. Chartoff, and W.M. Gelbart. 1995. Genetic characterization and cloning of mothers against dpp, a gene required for decapentaplegic function in Drosophila melanogaster. Genetics **139**:1347-1358.

Selbert, S., D.J. Bentley, D.W. Melton, D. Rannie, P. Lourenco, C.J. Watson, and A.R. Clarke. 1998. Efficient BLG-Cre mediated gene deletion in the mammary gland. Transgenic Res **7**:387-396.

Selfridge, J., A.M. Pow, J. McWhir, T.M. Magin, and D.W. Melton. 1992. Gene targeting using a mouse HPRT minigene/HPRT-deficient embryonic stem cell system: inactivation of the mouse ERCC-1 gene. Somat Cell Mol Genet **18**:325-336.

Shi, X., X. Yang, D. Chen, Z. Chang, and X. Cao. 1999. Smad1 interacts with homeobox DNA-binding proteins in bone morphogenetic protein signaling. J Biol Chem **274**:13711-13717.

Shi, Y., A. Hata, R.S. Lo, J. Massague, and N.P. Pavletich. 1997. A structural basis for mutational inactivation of the tumour suppressor Smad4 [see comments]. Nature **388**:87-93.

Shi, Y., Y.F. Wang, L. Jayaraman, H. Yang, J. Massague, and N.P. Pavletich. 1998. Crystal structure of a Smad MH1 domain bound to DNA: insights on DNA binding in TGF-beta signaling. Cell **94**:585-594.

Shigeta, T., M. Takagi, D. Delia, L. Chessa, S. Iwata, Y. Kanke, M. Asada, M. Eguchi, and S. Mizutani. 1999. Defective control of apoptosis and mitotic spindle checkpoint in heterozygous carriers of ATM mutations. Cancer Res **59**:2602-2607.

Shull, M.M., I. Ormsby, A.B. Kier, S. Pawlowski, R.J. Diebold, M. Yin, R. Allen, C. Sidman, G. Proetzel, and D. Calvin. 1992. Targeted disruption of the mouse transforming growth factor-beta 1 gene results in multifocal inflammatory disease. Nature **359**:693-699.

Sirard, C., J.L. de la Pompa, A. Elia, A. Itie, C. Mirtsos, A. Cheung, S. Hahn, A. Wakeham, L. Schwartz, S.E. Kern, J. Rossant, and T.W. Mak. 1998. The tumor suppressor gene Smad4/Dpc4 is required for gastrulation and later for anterior development of the mouse embryo. Genes Dev **12**:107-119.

Skowronski, J., D. Parks, and R. Mariani. 1993. Altered T cell activation and development in transgenic mice expressing the HIV-1 nef gene. EMBO J **12**:703-713.

Smith, A.G. and M.L. Hooper. 1987. Buffalo rat liver cells produce a diffusible activity which inhibits the differentiation of murine embryonal carcinoma and embryonic stem cells. Dev Biol **121**:1-9.

Smith, A.G., J.K. Heath, D.D. Donaldson, G.G. Wong, J. Moreau, M. Stahl, and D. Rogers. 1988. Inhibition of pluripotential embryonic stem cell differentiation by purified polypeptides. Nature **336**:688-690.

Smithies, O., R.G. Gregg, S.S. Boggs, M.A. Koralewski, and R.S. Kucherlapati. 1985. Insertion of DNA sequences into the human chromosomal beta-globin locus by homologous recombination. Nature **317**:230-234.

Song, C.Z., T.E. Siok, and T.D. Gelehrter. 1998. Smad4/DPC4 and Smad3 mediate transforming growth factor-beta (TGF-beta) signaling through direct binding to a novel TGF-beta-responsive element in the human plasminogen activator inhibitor-1 promoter. J Biol Chem **273**:29287-29290.

Spencer, F.A., F.M. Hoffmann, and W.M. Gelbart. 1982. Decapentaplegic: a gene complex affecting morphogenesis in Drosophila melanogaster. Cell **28**:451-461.

St-Onge, L., P.A. Furth, and P. Gruss. 1996. Temporal control of the Cre recombinase in transgenic mice by a tetracycline responsive promoter. Nucleic Acids Res **24**:3875-3877.

Stacey, A., J. Bateman, T. Choi, T. Mascara, W. Cole, and R. Jaenisch. 1988. Perinatal lethal osteogenesis imperfecta in transgenic mice bearing an engineered mutant pro-alpha 1(I) collagen gene. Nature **332**:131-136.

Sternberg, N., S. Austin, D. Hamilton, and M. Yarmolinsky. 1978. Analysis of bacteriophage P1 immunity by using lambda-P1 recombinants constructed in vitro. Proc Natl Acad Sci U S A **75**:5594-5598.

Sternberg, N. and D. Hamilton. 1981. Bacteriophage P1 site-specific recombination. I. Recombination between loxP sites. J Mol Biol **150**:467-486.

Sternberg, N., B. Sauer, R. Hoess, and K. Abremski. 1986. Bacteriophage P1 cre gene and its regulatory region. Evidence for multiple promoters and for regulation by DNA methylation. J Mol Biol **187**:197-212.

Strange, R., R.R. Friis, L.T. Bemis, and F.J. Geske. 1995. Programmed cell death during mammary gland involution. Methods Cell Biol **46**:355-368.

Stroschein, S.L., W. Wang, S. Zhou, Q. Zhou, and K. Luo. 1999. Negative feedback regulation of TGF-beta signaling by the SnoN oncoprotein [see comments]. Science **286**:771-774.

Stroschein, S.L., W. Wang, and K. Luo. 1999. Cooperative binding of Smad proteins to two adjacent DNA elements in the plasminogen activator inhibitor-1 promoter mediates transforming growth factor beta-induced smad-dependent transcriptional activation. J Biol Chem **274**:9431-9441.

Sun, Y., X. Liu, E.N. Eaton, W.S. Lane, H.F. Lodish, and R.A. Weinberg. 1999. Interaction of the Ski oncoprotein with Smad3 regulates TGF-beta signaling. Mol Cell **4**:499-509.

Sun, Y., X. Liu, E. Ng-Eaton, H.F. Lodish, and R.A. Weinberg. 1999. SnoN and Ski protooncoproteins are rapidly degraded in response to transforming growth factor beta signaling. Proc Natl Acad Sci U S A **96**:12442-12447.

Takeda, K., T. Kaisho, N. Yoshida, J. Takeda, T. Kishimoto, and S. Akira. 1998. Stat3 activation is responsible for IL-6-dependent T cell proliferation through preventing apoptosis: generation and characterization of T cell-specific Stat3-deficient mice. J Immunol **161**:4652-4660.

Tanaka, S. and T. Takeuchi. 1991. Hair pigmentation in transgenic mice. Ann N Y Acad Sci **642**:407-418.

Tanaka, T., K. Tanaka, S. Ogawa, M. Kurokawa, K. Mitani, Y. Yazaki, Y. Shibata, and H. Hirai. 1997. An acute myeloid leukemia gene, AML1, regulates transcriptional activation and hemopoietic myeloid cell differentiation antagonistically by two alternative spliced forms. Leukemia **11 Suppl 3**:299-302.

Tanaka, T., N. Yoshida, T. Kishimoto, and S. Akira. 1997. Defective adipocyte differentiation in mice lacking the C/EBPbeta and/or C/EBPdelta gene. EMBO J **16**:7432-7443.

Tanaka, Y., O. Williams, R. Tarazona, A. Wack, T. Norton, and D. Kioussis. 1997. In vitro positive selection of alpha beta TCR transgenic thymocytes by a conditionally immortalized cortical epithelial clone. Int Immunol **9**:381-393.

te, R., E.R. Maandag, and A. Berns. 1992. Highly efficient gene targeting in embryonic stem cells through homologous recombination with isogenic DNA constructs. Proc Natl Acad Sci U S A **89**:5128-5132.

ten Dijke, P., H. Yamashita, T.K. Sampath, A.H. Reddi, M. Estevez, D.L. Riddle, H. Ichijo, C.H. Heldin, and K. Miyazono. 1994. Identification of type I receptors for osteogenic protein-1 and bone morphogenetic protein-4. J Biol Chem **269**:16985-16988.

ten Dijke, P., P. Franzen, H. Yamashita, H. Ichijo, C.H. Heldin, and K. Miyazono. 1994. Serine/threonine kinase receptors. Prog Growth Factor Res **5**:55-72.

ten Dijke, P., K. Miyazono, and C.H. Heldin. 1996. Signaling via hetero-oligomeric complexes of type I and type II serine/threonine kinase receptors. Curr Opin Cell Biol **8**:139-145.

ten Dijke, P., K. Miyazono, and C.H. Heldin. 2000. Signaling inputs converge on nuclear effectors in TGF-beta signaling. Trends Biochem Sci **25**:64-70.

Teramoto, T., A. Kiss, and S.S. Thorgeirsson. 1998. Induction of p53 and Bax during TGF-beta 1 initiated apoptosis in rat liver epithelial cells. Biochem Biophys Res Commun **251**:56-60.

Terry, R.W., L. Kwee, H.S. Baldwin, and M.A. Labow. 1997. Cre-mediated generation of a VCAM-1 null allele in transgenic mice. Transgenic Res **6**:349-356.

Thomas, K.R. and M.R. Capecchi. 1987. Site-directed mutagenesis by gene targeting in mouse embryo-derived stem cells. Cell **51**:503-512.

Thompson, S., A.R. Clarke, A.M. Pow, M.L. Hooper, and D.W. Melton. 1989. Germ line transmission and expression of a corrected HPRT gene produced by gene targeting in embryonic stem cells. Cell **56**:313-321.

Thomsen, G.H. 1996. Xenopus mothers against decapentaplegic is an embryonic ventralizing agent that acts downstream of the BMP-2/4 receptor. Development **122**:2359-2366.

Topper, Y.J. and C.S. Freeman. 1980. Multiple hormone interactions in the developmental biology of the mammary gland. Physiol Rev **60**:1049-1106.

Tremml, G. and M. Bienz. 1992. Induction of labial expression in the Drosophila endoderm: response elements for dpp signalling and for autoregulation. Development **116**:447-456.

Tsukada, T., K. Eguchi, K. Migita, Y. Kawabe, A. Kawakami, N. Matsuoka, H. Takashima, A. Mizokami, and S. Nagataki. 1995. Transforming growth factor beta 1 induces apoptotic cell death in cultured human umbilical vein endothelial cells with down-regulated expression of bcl-2. Biochem Biophys Res Commun **210**:1076-1082.

Tsukazaki, T., T.A. Chiang, A.F. Davison, L. Attisano, and J.L. Wrana. 1998. SARA, a FYVE domain protein that recruits Smad2 to the TGFbeta receptor. Cell **95**:779-791.

Tsuneizumi, K., T. Nakayama, Y. Kamoshida, T.B. Kornberg, J.L. Christian, and T. Tabata. 1997. Daughters against dpp modulates dpp organizing activity in Drosophila wing development. Nature **389**:627-631.

Udy, G.B., R.P. Towers, R.G. Snell, R.J. Wilkins, S.H. Park, P.A. Ram, D.J. Waxman, and H.W. Davey. 1997. Requirement of STAT5b for sexual dimorphism of body growth rates and liver gene expression. Proc Natl Acad Sci U S A **94**:7239-7244.

Ulloa, L., J. Doody, and J. Massague. 1999. Inhibition of transforming growth factor-beta/SMAD signalling by the interferon-gamma/STAT pathway. Nature **397**:710-713.

Varlet, I., J. Collignon, and E.J. Robertson. 1997. nodal expression in the primitive endoderm is required for specification of the anterior axis during mouse gastrulation. Development **124**:1033-1044.

Vassalli, A., M.M. Matzuk, H.A. Gardner, K.F. Lee, and R. Jaenisch. 1994. Activin/inhibin beta B subunit gene disruption leads to defects in eyelid development and female reproduction. Genes Dev **8**:414-427.

Verschueren, K., J.E. Remacle, C. Collart, H. Kraft, B.S. Baker, P. Tylzanowski, L. Nelles, G. Wuytens, M.T. Su, R. Bodmer, J.C. Smith, and D. Huylebroeck. 1999. SIP1, a novel zinc finger/homeodomain repressor, interacts with Smad proteins and binds to 5'-CACCT sequences in candidate target genes. J Biol Chem **274**:20489-20498.

Vindelov, L.L., I.J. Christensen, and N.I. Nissen. 1983. A detergent-trypsin method for the preparation of nuclei for flow cytometric DNA analysis. Cytometry **3**:323-327.

Vivien, D. and J.L. Wrana. 1995. Ligand-induced recruitment and phosphorylation of reduced TGF-beta type I receptor. Exp Cell Res **221**:60-65.

Wagner, K.U., R.J. Wall, L. St-Onge, P. Gruss, A. Wynshaw-Boris, L. Garrett, M. Li, P.A. Furth, and L. Hennighausen. 1997. Cre-mediated gene deletion in the mammary gland. Nucleic Acids Res **25**:4323-4330.

Wagner, K.U., W.S. Young, X. Liu, E.I. Ginns, M. Li, P.A. Furth, and L. Hennighausen. 1997. Oxytocin and milk removal are required for post-partum mammary-gland development. Genes Funct **1**:233-244.

Wagner, T., Z. Swierczynska, C. Stankiewicz, H. Bujalska, and E. Kosowska. 1981. A new model of experimental amylcidosis? Z Rheumatol **40**:234-235.

Wahl, S.M., N. McCartney-Francis, and S.E. Mergenhagen. 1989. Inflammatory and immunomodulatory roles of TGF-beta. Immunol Today **10**:258-261.

Wahl, S.M. 1994. Transforming growth factor beta: the good, the bad, and the ugly. J Exp Med **180**:1587-1590.

Wang, T., P.K. Donahoe, and A.S. Zervos. 1994. Specific interaction of type I receptors of the TGF-beta family with the immunophilin FKBP-12. Science **265**:674-676.

Wang, T., B.Y. Li, P.D. Danielson, P.C. Shah, S. Rockwell, R.J. Lechleider, J. Martin, T. Manganaro, and P.K. Donahoe. 1996. The immunophilin FKBP12 functions as a common inhibitor of the TGF beta family type I receptors. Cell **86**:435-444.

Wang, Y., L.A. Krushel, and G.M. Edelman. 1996. Targeted DNA recombination in vivo using an adenovirus carrying the cre recombinase gene. Proc Natl Acad Sci U S A **93**:3932-3936.

Weber, M.S., P.L. Boyle, B.A. Corl, E.A. Wong, F.C. Gwazdauskas, and R.M. Akers. 1998. Expression of ovine insulin-like growth factor-1 (IGF-1) stimulates alveolar bud development in mammary glands of transgenic mice. Endocrine **8**:251-259.

Weinstein, M., X. Yang, C. Li, X. Xu, J. Gotay, and C.X. Deng. 1998. Failure of egg cylinder elongation and mesoderm induction in mouse embryos lacking the tumor suppressor smad2. Proc Natl Acad Sci U S A **95**:9378-9383.

Weis-Garcia, F. and J. Massague. 1996. Complementation between kinase-defective and activation-defective TGF-beta receptors reveals a novel form of receptor cooperativity essential for signaling. EMBO J **15**:276-289.

Weisberg, E., G.E. Winnier, X. Chen, C.L. Farnsworth, B.L. Hogan, and M. Whitman. 1998. A mouse homologue of FAST-1 transduces TGF beta superfamily signals and is expressed during early embryogenesis. Mech Dev **79**:17-27.

Wharton, K.A., R.P. Ray, and W.M. Gelbart. 1993. An activity gradient of decapentaplegic is necessary for the specification of dorsal pattern elements in the Drosophila embryo. Development **117**:807-822.

Whitelaw, C.B., S. Harris, M. McClenaghan, J.P. Simons, and A.J. Clark. 1992. Position-independent expression of the ovine beta-lactoglobulin gene in transgenic mice. Biochem J **286 (Pt 1)**:31-39.

Wieser, R., J.L. Wrana, and J. Massague. 1995. GS domain mutations that constitutively activate T beta R-I, the downstream signaling component in the TGF-beta receptor complex. EMBO J **14**:2199-2208.

Williams, R.L., D.J. Hilton, S. Pease, T.A. Willson, C.L. Stewart, D.P. Gearing, E.F. Wagner, D. Metcalf, N.A. Nicola, and N.M. Gough. 1988. Myeloid leukaemia inhibitory factor maintains the developmental potential of embryonic stem cells. Nature **336**:684-687.

Willis, S.A., C.M. Zimmerman, L.I. Li, and L.S. Mathews. 1996. Formation and activation by phosphorylation of activin receptor complexes. Mol Endocrinol **10**:367-379.

Wotton, D., R.S. Lo, S. Lee, and J. Massague. 1999. A Smad transcriptional corepressor. Cell **97**:29-39.

Wrana, J.L., L. Attisano, J. Carcamo, A. Zentella, J. Doody, M. Laiho, X.F. Wang, and J. Massague. 1992. TGF beta signals through a heteromeric protein kinase receptor complex. Cell **71**:1003-1014.

Wrana, J.L., H. Tran, L. Attisano, K. Arora, S.R. Childs, J. Massague, and M.B. O'Connor. 1994. Two distinct transmembrane serine/threonine kinases from Drosophila melanogaster form an activin receptor complex. Mol Cell Biol **14**:944-950.

Wrana, J.L. and L. Attisano. 1996. MAD-related proteins in TGF-beta signalling. Trends Genet **12**:493-496.

Wrana, J.L. 1998. TGF-beta receptors and signalling mechanisms. Miner Electrolyte Metab **24**:120-130.

Wrana, J.L. and L. Attisano. 2000. The Smad pathway. Cytokine Growth Factor Rev **11**:5-13.

Wu, G., Y.G. Chen, B. Ozdamar, C.A. Gyuricza, P.A. Chong, J.L. Wrana, J. Massague, and Y. Shi. 2000. Structural basis of Smad2 recognition by the Smad anchor for receptor activation. Science **287**:92-97.

Xu, P.X., J. Adams, H. Peters, M.C. Brown, S. Heaney, and R. Maas. 1999. Eya1-deficient mice lack ears and kidneys and show abnormal apoptosis of organ primordia. Nat Genet **23**:113-117.

Xu, X., S.G. Brodie, X. Yang, Y.H. Im, W.T. Parks, L. Chen, Y.X. Zhou, M. Weinstein, S.J. Kim, and C.X. Deng. 2000. Haploid loss of the tumor suppressor Smad4/Dpc4 initiates gastric polyposis and cancer in mice. Oncogene **19**:1868-1874.

Yagi, K., D. Goto, T. Hamamoto, S. Takenoshita, M. Kato, and K. Miyazono. 1999. Alternatively spliced variant of Smad2 lacking exon 3. Comparison with wild-type Smad2 and Smad3. J Biol Chem **274**:703-709.

Yamashita, H., P. ten Dijke, P. Franzen, K. Miyazono, and C.H. Heldin. 1994. Formation of hetero-oligomeric complexes of type I and type II receptors for transforming growth factor-beta. J Biol Chem **269**:20172-20178.

Yan, Q. and E.H. Sage. 1998. Transforming growth factor-beta1 induces apoptotic cell death in cultured retinal endothelial cells but not pericytes: association with decreased expression of p21waf1/cip1. J Cell Biochem **70**:70-83.

Yanagisawa, J., Y. Yanagi, Y. Masuhiro, M. Suzawa, M. Watanabe, K. Kashiwagi, T. Toriyabe, M. Kawabata, K. Miyazono, and S. Kato. 1999. Convergence of transforming growth factor-beta and vitamin D signaling pathways on SMAD transcriptional coactivators. Science **283**:1317-1321.

Yanagisawa, K., H. Osada, A. Masuda, M. Kondo, T. Saito, Y. Yatabe, K. Takagi, and T. Takahashi. 1998. Induction of apoptosis by Smad3 and down-regulation of Smad3 expression in response to TGF-beta in human normal lung epithelial cells. Oncogene **17**:1743-1747.

Yang, X., J.J. Letterio, R.J. Lechleider, L. Chen, R. Hayman, H. Gu, A.B. Roberts, and C. Deng. 1999. Targeted disruption of SMAD3 results in impaired mucosal immunity and diminished T cell responsiveness to TGF-beta. EMBO J **18**:1280-1291.

Yeo, C.Y., X. Chen, and M. Whitman. 1999. The role of FAST-1 and Smads in transcriptional regulation by activin during early Xenopus embryogenesis. J Biol Chem **274**:26584-26590.

Yingling, J.M., M.B. Datto, C. Wong, J.P. Frederick, N.T. Liberati, and X.F. Wang. 1997. Tumor suppressor Smad4 is a transforming growth factor beta-inducible DNA binding protein. Mol Cell Biol **17**:7019-7028.

Young, W.S., E. Shepard, A.C. DeVries, A. Zimmer, M.E. LaMarca, E.I. Ginns, J. Amico, R.J. Nelson, L. Hennighausen, and K.U. Wagner. 1998. Targeted reduction of oxytocin expression provides insights into its physiological roles. Adv Exp Med Biol **449**:231-240.

Zavizion, B., C.W. Heegard, J. White, F. Cheli, and I. Politis. 1996. Synthesis of plasminogen activator inhibitor 1 by bovine mammary epithelial and myoepithelial cell lines. J Dairy Res **63**:451-458.

Zawel, L., J.L. Dai, P. Buckhaults, S. Zhou, K.W. Kinzler, B. Vogelstein, and S.E. Kern. 1998. Human Smad3 and Smad4 are sequence-specific transcription activators. Mol Cell **1**:611-617.

Zhang, G., V. Gurtu, S.R. Kain, and G. Yan. 1997. Early detection of apoptosis using a fluorescent conjugate of annexin V. Biotechniques **23**:525-531.

Zhang, Y., X. Feng, R. We, and R. Derynck. 1996. Receptor-associated Mad homologues synergize as effectors of the TGF-beta response. Nature **383**:168-172.

Zhang, Y., X. Feng, R. We, and R. Derynck. 1996. Receptor-associated Mad homologues synergize as effectors of the TGF-beta response. Nature **383**:168-172.

Zhang, Y., T. Musci, and R. Derynck. 1997. The tumor suppressor Smad4/DPC 4 as a central mediator of Smad function. Curr Biol **7**:270-276.

Zhang, Y., X.H. Feng, and R. Derynck. 1998. Smad3 and Smad4 cooperate with c-Jun/c-Fos to mediate TGF-beta-induced transcription [published erratum appears in Nature 1998 Dec 3;396(6710):491]. Nature **394**:909-913.

Zhang, Y. and R. Derynck. 1999. Regulation of Smad signalling by protein associations and signalling crosstalk. Trends Cell Biol **9**:274-279.

Zhou, S., L. Zawel, C. Lengauer, K.W. Kinzler, and B. Vogelstein. 1998. Characterization of human FAST-1, a TGF beta and activin signal transducer. Mol Cell **2**:121-127.

Zhou, X., H. Sasaki, L. Lowe, B.L. Hogan, and M.R. Kuehn. 1993. Nodal is a novel TGF-beta-like gene expressed in the mouse node during gastrulation. Nature **361**:543-547.

Zhu, H., P. Kavsak, S. Abdollah, J.L. Wrana, and G.H. Thomsen. 1999. A SMAD ubiquitin ligase targets the BMP pathway and affects embryonic pattern formation. Nature **400**:687-693.

Zhu, Y., J.A. Richardson, L.F. Parada, and J.M. Graff. 1998. Smad3 mutant mice develop metastatic colorectal cancer. Cell **94**:703-714.

Zimmerman, L.B., J.M. De Jesus-Escobar, and R.M. Harland. 1996. The Spemann organizer signal noggin binds and inactivates bone morphogenetic protein 4. Cell **86**:599-606.

Appendix A – solutions and distributors

L-Broth

Bactotryptone (*DIFCO*)	10g/L
Yeast Extract (*DIFCO*)	5g/L
NaCl (*Sigma*)	10g/L

In de-ionised H_2O, autoclaved

Supplemented with 50µg/ml ampicillin where stated

Supplemented with 1.2% w/v Bactoagar for plates

Solution P1

Tris-HCl, pH 8.0	50mM
EDTA (*Sigma*)	10mM
Rnase A (*Sigma*)	100µg/ml

Should be stored at 4°C after addition of Rnase A

Solution P2

NaOH (*Fisher Scientific*)	200mM
Sodium Dodecyl Sulphate (SDS) (*ICN, US*)	1%

Should be stored at room temperature

Solution P3

Potassium acetate, pH 5.5 (*Fisher*)	3.0 M

Should be stored at room temperature

Buffer QBT

NaCl (*Sigma*)	750 mM
MOPS, pH 7.0 (*Sigma*)	50mM
Isopropanol (*Fisher*)	15% v/v
Triton X-100 (*BioRad, US*)	0.15% v/v

Should be stored at room temperature.

Buffer QC

NaCl (*Sigma*)	1.0M
MOPS, pH 7.0 (*Sigma*)	50mM
Isopropanol (*Fisher*)	15 %

Should be stored at room temperature.

Buffer QF

NaCl (*Sigma*)	1.25 M
MOPS, pH 7.0 (*Sigma*)	50mM
Isopropanol (*Fisher*)	15%

Should be stored at room temperature.

10 x TBE Buffer:

TRIZMA base (*Sigma*)	108g
Boric Acid (*Sigma*)	55g
0.5M EDTA pH 8.0 (*Fisher*)	40mls

Make up to 1L with double distilled water

10 x Loading Buffer

Bromophenol blue (*Fisher*)	0.25%
Glycerol (*Sigma*)	25%

Make up in double distilled water

Sequencing Gel Solution

Urea (*Sigma*)	250g
19:1, 40% acrylamide stock (*Severn Biotech Ltd. UK*)	75mls
10 x TBE	50mls
Double distilled water	175mls

Tail Lysis Buffer

TRIZMA base (*Sigma*)	100mM, pH 8.5
EDTA (*Fisher*)	5mM
SDS (*ICN*)	0.2%
NaCl (*Sigma*)	200mM

LacZ Stain

NaH_2PO_4 (*Fisher*)	20mM
Na_2HPO_4 (*Fisher*)	80mM
$K_3Fe(CN)_6$ (*Fisher*)	1.3mM
$K_4Fe(CN)_6$ (*Fisher*)	3mM
X-Gal (*Promega*)	1mg/ml

(5-Bromo-4-chloro-3-I indolyl-beta-D-galactopyranoside)

Binding Buffer

1M HEPES pH 7.4 (*BDH*)	0.1mls
1M NaCl (*Sigma*)	1.4mls
1M $CaCl_2.6H_20$ (*Sigma*)	25µl
Double distilled water	8.5mls
Propidium Iodide	50µg/ml in PBS

Citrate Buffer

Sucrose (*Fisher*)	85.5g
Trisodium citrate (*Fisher*)	11.76g
Double distilled water	800ml
DMSO (*Sigma*)	50mls

pH to 7.6 with 5M HCl

Stock Solution

Tri-sodium citrate (*Sigma*)	2g
TRIZMA base (*Sigma*)	121 mg
Spermidine tetrahydrochloride (*Sigma*)	1044mg
Nonidet P40 (*Sigma*)	2mls
Double distilled water	2L

Adjust pH to 7.6 with 5M HCl

Solution A

Trypsin (*Sigma*) in 500mls Stock Solution, store frozen	15mg

Solution B

Trypsin inhibitor (*Sigma*)	250mg
Ribonuclease A (*Sigma*)	50mg

In 500mls Stock Solution, store frozen

Solution C

Propidium iodide	208mg
Spermidine tetrachloride in 500mls (*Sigma*)	500mg

In 500mls Stock Solution, store frozen